

١١٤٠٤٢

مقدمة في الاقتصاد القياسي

المبادئ والتطبيقات

تأليف

والاس أوتس

Wallace E. Oates

هارى كلجيان

Harry H. Kelejian

ترجمة

د. المرسى السيد حجازى و د. عبدالقادر محمد عطية

أستاذ مشارك، قسم الاقتصاد، كلية الاقتصاد والإدارة

جامعة الملك سعود (سابقاً)

مراجعة علمية

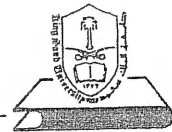
د. حمد بن سليمان البازعي

أستاذ مساعد، قسم الاقتصاد، كلية الاقتصاد والإدارة

جامعة الملك سعود، فرع القصيم

النشر العلمي والمطابع - جامعة الملك سعود

ص.ب ٦٨٩٥٣ - الرياض ١١٥٣٧ - المملكة العربية السعودية



ح جامعة الملك سعود، ١٤٢٢هـ

الطبعة الأولى: ١٤٢٢هـ (٢٠٠١م)

الترجمة العربية للطبعة الثالثة من كتاب:

Introduction to Econometrics: Principle and Applications

© 1989, Harry H. Kelejian and Wallace E. Oates, 3rd edition.

Published By: Harper & Row, Publishers, Inc.

فهرسة مكتبة الملك فهد الوطنية أثناء النشر

كلجيان، هاري

مقدمة في الاقتصاد القياسي: المبادئ والتطبيقات /

هاري كلجيان، والاس أوتس، ترجمة المرسي السيد حجازي،

عبدالقادر محمد عطية، ط ١. الرياض

٥٤٦ ص، ١٧ سم × ٢٤ سم

ردمك ٩٩٦٠-٠٥-٩٨٣-٩

١- الاقتصاد القياسي أ- أوتس، والاس (م. مشارك)

ب- حجازي، المرسي السيد (مترجم) ج- عطية، عبدالقادر

محمد (مترجم) د- العنوان

٢٠ / ١١٤٦

ديوي ٣٣٩، ٣

رقم الإيداع ٢٠ / ١١٤٦

تم تحكيم الكتاب بواسطة لجنة متخصصة شكلها المجلس العلمي بالجامعة. وقد وافق المجلس العلمي على نشره، بعد اطلاعه على تقارير المحكمين، في اجتماعه الثامن للعام الدراسي ١٤١٦/١٤١٧هـ المعقود بتاريخ ٩/٨/١٤١٦هـ الموافق ٣١/١٢/١٩٩٥م.

النشر العلمي والمطابع ١٤٢٢هـ



تقديم

الحمد لله رب العالمين ، والصلاة والسلام على نبينا محمد وعلى آله وصحبه أجمعين . . . وبعد ،

فقد انتهجت كلية الاقتصاد والإدارة بجامعة الملك سعود فرع القصيم منذ إنشائها في العام الجامعي (١٤٠١/١٤٠٢ هـ الموافق ١٩٨١/١٩٨٢ م) نهجاً خاصاً قام على ترجمة الكتب الدراسية والمرجعية في مجالات تخصصاتها وتعريبها . وهذا التوجه حتمته الحاجة لإعداد جيل من الخريجين والخريجات على قدر رفيع من الكفاءة العصرية والأداء المتخصص في عالم يتسم بالتطور السريع المستمر . وقد كان للكفاءات المتخصصة البارزة التي تستقطبها الكلية الأثر الإيجابي الفعال في جهود الترجمة والتعريب مما ساعد على انتشار استخدام الكتب المترجمة والمعرّبة داخل المملكة وخارجها ، حيث تدرس معظم هذه الكتب في الكليات والأقسام العلمية ذات العلاقة سواء كان ذلك كتباً دراسية لمقررات هذه الكتب أو مراجع مساعدة . وقد زاد هذا مسئولية الكلية تجاه ترجمة الكتب العلمية وتعريبها ، الأمر الذي ترحب به الكلية دائماً . فوضع كتب قيمة بين يدي القارئ العربي أمر حثنا عليه ديننا الحنيف . ولن يؤدي مثل هذا الكتاب إلى تحسين المعرفة لدى الطالب الدارس ، فقط ، بل سيفيد ، أيضاً ، في تطوير المادة العلمية التي تحتويها الكتب المؤلفة بالعربية في موضوع الكتاب المترجم أو المعرب نفسه .

وفي إطار نشاط الكلية في مجال الترجمة والتعريب ، وبعد دراسة متأنية للكتب في مجال الاقتصاد القياسي Econometrics ، وقع اختيار الكلية (مثلة في قسم الاقتصاد بها) على كتاب «مقدمة في الاقتصاد القياسي : المبادئ والتطبيقات» (الطبعة الثالثة -

١٩٨٩م) لمؤلفيه هاري كلجيان Harry H. Kelejian وولاس أوتس Wallace Oates الأستاذان بجامعة مرييلاند بالولايات المتحدة الأمريكية . ويشتمل الكتاب على أحدث التطورات التي طرأت على مجال القياس الكمي للعلاقة بين المتغيرات الاقتصادية . ولا يفوتني أن انوه بالجهد العلمي المشكور الذي بذله كل من الدكتور المرسي السيد أحمد حجازي الأستاذ المشارك بقسم الاقتصاد بالكلية (سابقاً) والدكتور عبد القادر محمد عبد القادر عطية الأستاذ المشارك بقسم الاقتصاد بالكلية (سابقاً) في ترجمة هذا الكتاب إلى اللغة العربية ، داعياً الله أن يجزيهما خير الجزاء على عملهما هذا . ونسأل الله أن يبارك في جهودنا وأن يجعل جميع أعمالنا خالصة لوجهه الكريم إنه سميع مجيب

د. حمد بن سليمان البازعي

عميد كلية الاقتصاد والإدارة بالنيابة (سابقاً)

مقدمة الترجمة العربية

الحمد لله رب العالمين والصلاة والسلام على رسول الله محمد بن عبد الله ، وعلى آله وصحبه أجمعين وبعد ،

ترجع فكرة ترجمة هذا الكتاب إلى العام الدراسي ١٤١٢/١٤١٣ هـ حينما قرر مجلس قسم الاقتصاد بالكلية البحث عن كتاب في مجال الاقتصاد القياسي ليكون مرجعاً دراسياً لمقرر ٤٣٣ قصد . بدأ البحث عن طريق مراسلة مجموعة من الجامعات الأمريكية لمعرفة كتب الاقتصاد القياسي التي يدرسها طلبة المستوى الجامعي الأول بها . وبعد الاستجابة الجيدة من تلك الجامعات ، استقر الأمر في النهاية على المفاضلة بين أربعة من الكتب المنتشرة ، في غالبيتها ، في مجال الاقتصاد القياسي . بعد أن حصل القسم على نسخ من هذه الكتب الأربعة ، قام أعضاء القسم بالاطلاع عليها وإعداد تقارير عنها ناقشها بعد ذلك مجلس قسم الاقتصاد في جلسة طويلة ، واستقر الأمر في النهاية على ترجمة هذا الكتاب .

يتميز هذا الكتاب بسهولة معالجة الموضوعات واستخدام الأساليب الرياضية الأولية التي يستطيع القارئ العادي ، فضلاً عن طالب المستوى الجامعي الأول ، استيعابها دون مواجهة صعوبة كبيرة . يشتمل الكتاب ، أيضاً ، على مقدمة إحصائية جيدة تلائم موضوعاته ، كما يحتوي كل فصل من فصوله على إطار نظري للموضوع تتلوه بعض الأمثلة التطبيقية الاقتصادية التي تثبت فهم الموضوع ، ثم يأتي ملحق أو أكثر في كل فصل لإثبات النظريات والعلاقات الرياضية المعقدة التي وردت به (يمكن للقارئ العادي إهمالها) ، ثم ينتهي الفصل بمجموعة من الأسئلة تساعد على الفهم

الأعمق للموضوعات النظرية والعملية الواردة به . وأخيراً، يشتمل الكتاب في آخر أجزائه على إجابات للأسئلة التي وردت في نهايات الفصول الثمانية مما يعطي فرصة جيدة للقارئ لمراجعة مدى استيعابه لموضوعاته .

يحتوي الكتاب على ثمانية فصول يتناول الأول منها مقدمة تحتوي على مراجعة عامة للمفاهيم الإحصائية، بينما يناقش الفصل الثاني نموذج الانحدار البسيط (ذي المتغيرين)، فيتعرض لقياس العلاقة الإحصائية بين متغيرين وتوضيح افتراضات النموذج ثم يقوم بتقدير معادلة الانحدار وبيان خصائص المقدرات الناتجة، وقياس القوة التفسيرية للنموذج، ثم ينتهي الفصل بمثال تطبيقي؛ تقدير دالة التكاليف . أما الفصل الثالث فيتعرض لتطبيقات نموذج الانحدار حيث يناقش اختبار الفرضيات وفترات الثقة، والشكل الدالي للعلاقات الاقتصادية، واستخدام المتغيرات المبطة، والتنبؤ، ثم يعطي مثلاً تطبيقياً؛ تقدير دالة الطلب .

يتناول الفصل الرابع نموذج الانحدار المتعدد، فيناقش خصائص المقدرات ومعامل التحديد ومثالين تطبيين أحدهما لدالة الاستهلاك والآخر في مجال الضرائب . ويتناول الفصل الخامس طرقاً أخرى في تحليل الانحدار المتعدد، حيث يتعرض لموضوعات العلاقات المبطة واستخدام المتغيرات الصورية والأشكال الدالية المختلفة مع إعطاء مثال للطلب على النقود .

يأتي بعد ذلك الفصل السادس ليعالج أهم مشكلات الانحدار وبيان نتائج كل مشكلة على خواص المقدرات، وهي مشكلات تعدد العلاقات الخطية، والارتباط الذاتي، اختلاف التباين واختيار المتغيرات . ويناقش الفصل السابع موضوع نظم المعادلات؛ فيتعرض لقضايا تحيز المعادلات الآنية، وطريقة المربعات الصغرى ذات المرحلتين (م ص م) ومشكلة تمييز المعادلات، كما يعطي مثالين تطبيين هما نموذج للطلب والعرض وآخر للمالية العامة المحلية . ويتناول الفصل الثامن والآخر نموذج المعادلات الآنية غير الخطية حيث يعالج موضوعات إطار التحليل، ومشكلة التمييز، واستخدام (م ص م) ويعطي مثالا من الاقتصاد الكلي .

أما مسئولية ترجمة الكتاب ومراجعة تجارب الطبع وتوحيد المصطلحات العلمية به، وإضافة المصطلحات وكشاف الموضوعات في نهايته فتقع على عاتق

الدكتور المرسي السيد حجازي . بينما ساهم الدكتور عبدالقادر محمد عطية بترجمة الفصل الثاني ومقدمة طبعات الكتاب الثلاثة . وقام الدكتور حمد سليمان البازعي بالمراجعة العلمية للكتاب .

ومن أجل الوفاء ببعض الحقوق لأصحابها ، يود المترجمان أن يتقدما بخالص الشكر والتقدير إلى الدكتور حمد بن سليمان البازعي عميد الكلية بالنيابة على تشجيعه المتواصل ومشاركته الفعالة ، من خلال المراجعة ، في إخراج هذا العمل إلى حيز الوجود ، كما نشكر كلا من الدكتور ولاس أوتس وهاري كلجيان لترحيبهما وموافقتهما على نشر هذا الكتاب باللغة العربية . ونشكر مركز البحوث بكلية الاقتصاد والإدارة على المساهمة الكبيرة والمشاركة في تحمل مسؤولية طباعة هذا الكتاب وإخراجه . ولا يفوتنا بهذه المناسبة أن نشكر الأخوة العاملين في سكرتارية الكلية على جهدهم الوافر المشكور في إدخال الأصول الأولى من الترجمة وتصحيحها على جهاز الحاسب الآلي ، ونشكر مركز الترجمة والنشر بجامعة الملك سعود على تحكيم هذا الكتاب ونشره ، كما نشكر جميع الأخوة الزملاء أعضاء مجلس قسم الاقتصاد بالكلية على دعمهم ومساندتهم لإبراز هذا الكتاب إلى حيز الوجود .

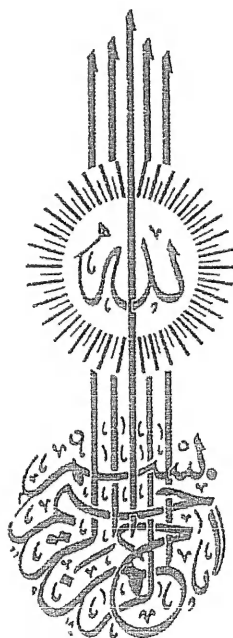
والله نسأل أن يكون هذا الكتاب إضافة طيبة ورصيداً علمياً جديداً للمكتبة الاقتصادية العربية وأن ينفع به وأن يجعله في ميزان أعمالنا الصالحة يوم القيامة ، والله ولي التوفيق .

المترجمان

د. المرسي السيد حجازي

د. عبد القادر محمد عطية

و ٢٠٧
عن ٤٥



مقدمة الطبعة الثالثة

سعيدنا في هذه الطبعة إلى سد ما قد رآه بعض القراء من ثغرات في الطبعتين السابقتين، وعلى وجه التحديد، فإن هذه الطبعة تحتوي على خمس إضافات :

١ - توسعة الملحق B في الفصل الأول الذي يقدم مراجعة للمفاهيم الإحصائية الأساسية. وتقدم هذه التوسعة مفهوم «دالة الكثافة المشتركة» وتطور عدد من النتائج المترتبة عليها التي يستخدم بعضها لاحقاً في متن الكتاب.

٢ - مناقشة موسعة عن قياس المقدرة التفسيرية لنموذج الانحدار. وهي تعد أحد مقاييس جودة التوفيق وأداة للاختيار من بين نماذج الانحدار المختلفة. ولهذا الغرض، نناقش عيوب مقياس R^2 ، ثم نقدم مقياساً بديلاً يسمى معامل التحديد المعدل، (أو إحصائية \bar{R}^2). ويظهر هذا المقياس في عدد من برامج الحاسوب. وسوف نناقش خصائصه، ثم نوضح علاقته بالإحصائية R^2 . وسوف نناقش، أيضاً، العلاقة بين قضايا اختيار النموذج واختبار الفروض. وفي هذا الإطار، نحذر من الاستخدام المبالغ فيه لـ \bar{R}^2 أو مقاييس جودة التوفيق الأخرى بغرض اختيار النماذج.

٣ - مناقشة مشكلات الارتباط الذاتي في حالة وجود متغير تابع مبطاً، وحيث إن اختبار دربن - واتسون Durbin-Watson test المتعارف عليه غير صالح للاستخدام في مثل هذه الحالات، فسوف نناقش اختبارين بديلين.

٤ - مناقشة مفهوم الاستقرار، وتظهر مشكلات الاستقرار في حالة نماذج الانحدار التي تحتوي على متغيرات تابعة مبطأة.

٥ - معالجة موسعة لمشكلة اختلاف تباين الخطأ العشوائي التي تشتمل على

اختبار جولدفيلد-كواندت Goldfeld-Quandt test . ويعد هذا الاختبار جذاباً ومباشراً في حالات معينة سوف نناقشها فيما بعد .
إننا ممتنون لقيام انجمار روشا بمراجعة مسودّات بعض الموضوعات الجديدة في هذه الطبعة وإبداء الملاحظات بشأنها ، وبالطبع ، لا يعد هذا تنصلاً من مبدأ المسؤولية المعتاد .

مقدمة الطبعة الثانية

هدفنا في هذه الطبعة ذو ثلاثة أبعاد : أولها زيادة عدد التطبيقات والتوضيحات الخاصة بالطرق القياسية المقدمة في هذا الكتاب ونطاقها، وثانيها توسيع التحليل ليشمل تقدير النماذج غير الخطية، وثالثها تصحيح معالجة بعض القضايا النظرية التي وردت في الطبعة الأولى وتوضيحها. ونتيجة لذلك، فلقد تضمنت الطبعة الثانية عددًا من الأمثلة الجديدة عن تحليل الارتباط والانحدار، كما أضفنا فصلاً جديداً هو الفصل الثامن عن تقدير النماذج غير الخطية.

أما الذين ألفوا الطبعة الأولى فسوف يجدون عددًا من التوضيحات الجديدة المتنوعة التي توضح للطالب كيفية القيام بعمل حسابات فعلية للمعلومات المقدرة. وقد أخذت أمثلة أخرى من أدبيات الاقتصاد لتوضح كيفية استخدام تحليل الانحدار، في الواقع، في تقدير معلمات مهمة، وفي اختبار بعض الفرضيات الأساسية في كل من الاقتصاد الجزئي والاقتصاد الكلي. وتتضمن الأمثلة الجديدة تحليل الارتباط البسيط، تقدير منحنيات الطلب على السلع الحقيقية وعلى الأرصدة النقدية، تقدير دالة التكاليف، دراسة حالة لمشكلة اختلاف التباين التي تنجم عن التجميع. وأملنا أن تساعد هذه الأمثلة الإضافية الطالب على فهم أفضل لتكوين النماذج القياسية، واستخدام البيانات الواقعية وتفسير النتائج.

وفي الواقع العملي، فإن معظم النماذج القياسية غير خطية. وبالرغم من ذلك، فإن معظم مراجع مرحلة البكالوريوس ومابعدھا تركّز، فقط، على النماذج الخطية تأسيساً على افتراض ظاهري هو صعوبة معالجة النماذج غير الخطية. ونحن مقتنعون بأن الأمر ليس كذلك، ولذا، فقد قدمنا في الفصل الثامن مناقشة موسعة للنماذج غير

الخطية على مستوى مرحلة البكالوريوس . وتأسيساً على المادة المعروضة في الكتاب من قبل ، فإن الفصل الجديد يدلف ليناقدش مشاكل التمييز والتقدير ، واختبار الفرضيات في النماذج غير الخطية ، كما يقدم تطبيقاً لهذه الفنون على مشكلة اقتصادية محددة . إن هدفنا الأساسي من هذه الطبعة لم يتغير عنه في الطبعة الأولى ، فنحن نسعى إلى تقديم تشكيلة من الطرق القياسية التي تتطلب فقط مهارات أولية في الرياضيات والإحصاء . ونحيل القارئ إلى مقدمة الطبعة الأولى المرفقة للوقوف على وصف طريقتنا .

ونود أن نعبر عن امتناننا لستيفن جولد فيلد ، ورونالد أوكساكا وريتشارد كواندت ، لتعليقاتهم البناءة على مسودات الفصل الثامن الجديد . كما ندين بالشكر للراحل رونالد فيشر وللناشر أوليفر وبويد ادنبره لسماحهم بإعادة طباعة جدول رقم ٢ من كتابهم : الطرق الإحصائية للعاملين في مجال البحوث .

مقدمة الطبعة الأولى

لقد اشتمل التطور الحديث في علم الاقتصاد على تقدم ملحوظ في الطرق القياسية وتطبيقاتها في التحليل الاقتصادي . وبعد أن كان الاقتصاد القياسي مقصوراً على قلة مختارة ، أصبح الآن مكوناً أساسياً في تدريب جميع دارسي الاقتصاد .

وبالرغم من تزايد انتشار استخدام التحليل القياسي ، فإن معظم المراجع التي تقدم مدى معقول من النتائج مازالت تتطلب مستوى مرتفعاً جداً من التأهيل الرياضي الذي يفوق بكثير ما يمتلكه عديد من الطلاب . وهدفنا في هذا المرجع هو تقديم مادة علمية موسعة تحتاج إلى متطلبات رياضية متواضعة معقولة من الطالب . وبتحديد أكثر ، فإن هذا الكتاب لا يستخدم حساب التكامل والتفاضل أو جبر المصفوفات . فنحن نفترض مستوى من المعرفة الرياضية يعادل تقريباً جبر الصف الثاني الثانوي * . وبالرغم من أن العرض يعتمد على طرق رياضية أولية ، فإن المادة المقدمة في هذا الكتاب تناظر المادة المقدمة في مقرر نمطي للاقتصاد القياسي على مستوى الدراسات العليا . فالموضوعات التي عولجت ، على سبيل المثال ، هي ، تقريباً ، المقدمة في كتاب جولد برجر Goldberger A.S. النظرية القياسية (Wiley 1964) ، Econometric Theory ، وكتاب جونستون J. Johnston « طرق قياسية » الطبعة الثانية Econometric Methods (McGraw-Hill 1972) .

وتشتق النتائج الأساسية في هذا الكتاب باستخدام طريقة المتغير المساعد ،

(*) يوجد في ملحق الفصل الأول بعض النتائج المرتبطة بصيغ الجمع المهمة والتي قد يحتاجها الطالب غير المتمرس عليها .

وتفوق هذه الطريقة طريقة المربعات الصغرى الأكثر استخدامًا بميزتين، الأولى أنها لا تتطلب حساب التفاضل والتكامل، والثانية أنها تسمح للطالب أن يرى، بوضوح، الدور الذي يؤديه كل افتراض عند إجراء التقدير. فعلى سبيل المثال، أوضحنا التناظر بين المعدلات الطبيعية والافتراضات الأساسية لنموذج الانحدار. وفي هذا الصدد، تم التركيز، خاصة على الإجراء الذي بمقتضاه يترجم كل افتراض معطى للمعادلة الطبيعية المناظرة. وهذه الطريقة مفيدة للغاية، حيث يمكن للطالب، فيما بعد، أن يرى، مباشرة، النتائج المترتبة على اختلال افتراض معين، كما يمكنه أن يفهم فهمًا أفضل الطرق المستخدمة لتعديل إجراء التقدير. ونستخدم، أيضًا، طريقة المتغير المساعد عبر الكتاب، وهذا يسمح بمعالجة موحدة للارتباط الذاتي، وتعدد العلاقات الخطية، واختلاف التباين، ومشكلات النظم وما إلى ذلك، حيث إنها - جميعًا - تعالج بالطريقة نفسها.

ولقد جرى التأكيد في هذا المرجع على ما قد يمكن تسميته بالحدسية «الفطنة»، بمعنى أننا لا نكتفي بذكر النتائج فحسب، وإنما تشتق بطريقة حدسية مع محاولة ترك أقل ما يمكن من النهايات غير المحددة. وبالرغم من أننا نعرض النتائج والحالات المعيارية، إلا أننا نركز أولاً على الإجراء الذي يحصل بمقتضاه على هذه النتائج، وثانيًا على تطبيقات هذه النتائج على مشكلات تقدير فعلية. وبعد تقديم كل طريقة جديدة، نوضح استخدامها بأثلة رقمية ودراسات فعلية من أدبيات الاقتصاد، ونتيجة لذلك، فإن الطالب، بعمله من خلال هذا المرجع، سوف يخرج بإدراك جيد عن كيفية عمل الأشياء وأسباب ذلك.

ولقد قدم عدد من التمارين في نهاية كل فصل، كما قُدمت إجابات لكل التمارين في نهاية الكتاب. وينصح الطالب المجد بحل هذه التمارين حيث إنها ذات صلة بكل من التطبيق العملي للنتائج المعروضة في الكتاب وبمعالجة المفاهيم ذات العلاقة وفهمها. بالإضافة إلى ذلك، قمنا بتقديم عدد من التمارين التي يعرض فيها النموذج في صورة لفظية ثم يسأل الطالب أن يقوم بصياغته في صورة نموذج انحدار. وسوف تعطي هذه التمارين الطالب فهمًا أفضل لبعض الصعوبات التي تكتنف صياغة نموذج اقتصادي.

لقد كتب هذا الكتاب ليستعمل مرجعاً في الاقتصاد القياسي يدرس على مدى فصل دراسي واحد في مرحلة البكالوريوس أو الماجستير . ونقصد بمستوى الماجستير هنا مقررات الدراسات العليا في الاقتصاد القياسي الموجهة للطلبة الذين لا تتوافر لديهم خلفية رياضية أو إحصائية تمكنهم من متابعة الكتابات المتقدمة في هذا العلم .

والمطلب السابق لهذا الكتاب هو ، تقريباً ، الثلثان الأولان لفصل دراسي في مبادئ الإحصاء . فعلى سبيل المثال ، تعدّ المادة المقدمة في الفصول الخمسة الأولى من كتاب W.C. Gunther مفاهيم الاستدلال الإحصائي Concepts of Statistical Inference أكثر من كافية . وأي معلومات إضافية يحتاج إليها استخدام هذا الكتاب (بالإضافة إلى المراجعة المختصرة للمفاهيم الإحصائية الأساسية) توجد في ملحق بالفصل الأول . وقد يعطي وصف مختصر لتطور هذا المخطوط إدراكاً أحسن لاستخداماته المحتملة . فلقد كانت نقطة انطلاق هذا الكتاب مجموعة مذكرات كلجيان Kelejian في مادة الاقتصاد القياسي المقررة على طلبة البكالوريوس بجامعة برنستون . وكانت هذه المذكرات توزع على الطلبة في صورة منسوخة على الآلة الكاتبة . وشجعت ردود فعل الطلبة في جامعة برنستون وفي أماكن أخرى على تأليف هذا الكتاب .

ولقد استخدمت المذكرات سالفة الذكر وكذلك مسودات مبكرة من فصول هذا المخطوط في تدريس مقرر لغير المتخصصين في الاقتصاد القياسي من طلبة الدراسات العليا بجامعة نيويورك ، كما استخدمت في تدريس مقرر للأساليب الكمية في برنامج الماجستير للشؤون العامة بمدرسة وودرو ويلسون Woodrow Wilson School في برنستون . وهذه هي أنواع المقررات التي نشعر بملاءمة هذا الكتاب لها . بالإضافة إلى ذلك ، وجد عدد من الطلبة ذوي المعرفة الأكثر تقدماً في مجال الاقتصاد القياسي أن هذا المخطوط يساعد في فهم أعمق لنتائج توصلوا إليها في مقررات أخرى أكثر تعقيداً .

ويعكس التعاون الخاص في مجال هذا الكتاب الهدف منه ، فيعد هاري كلجيان الاقتصاد القياسي مجال تخصصه الأساسي . حيث إن جهوده في البحث والتدريس كانت مركزة أساساً في هذا المجال . أما ولاس أوتس فاهتماماته الأساسية تنصب على مشكلات المالية الحكومية ، وعلاقته بالاقتصاد القياسي هي ، أساساً ، بوصفه

عمارساً ينصب اهتمامه على التحليل الكمي للمشكلات الاقتصادية الفعلية . وكان الأمل أن يؤدي هذا المزيج من الاهتمامات إلى تأليف كتاب يتعمق في مجال الاقتصاد القياسي وفي الوقت نفسه يكون مقروءاً لدى الطلبة الذين لم يسبق لهم دراسة الاقتصاد القياسي .

ونود - ختاماً - أن نعبر عن امتناننا لكل من قدّم مساعدة أو اقتراحات قيمة على مسودات الكتاب ومنهم شارلس بينشن ، ولاري هيرش ، وويليام لورانس ، وروبرت بلوتيناك ، وريتشارد كوانت ، وف . ساندراجان وايراسوهن ، وبالطبع لا يتحمل أحد منهم مسؤولية أي عيب مازال موجوداً في الكتاب . وبالإضافة إلى ذلك ، فنحن مدينون لراعي حقوق التأليف الخاصة بالراحل فيشر والناشر أوليفر وبويد أدنبره لسماحهم بإعادة طباعة جدول رقم ٢ من كتابهم «الطرق الإحصائية للباحثين» . أخيراً ، نود التعبير عن العرفات بالجميل للسيدة بيتي كامبسنسكي لطباعتها المتمرسة التي تمت في الغالب ، تحت ظروف صعبة .

المحتويات

الصفحة	
هـ	تقديم
ز	مقدمة الترجمة العربية
ك	مقدمة الطبعة الثالثة
م	مقدمة الطبعة الثانية
س	مقدمة الطبعة الأولى
	الفصل الأول : مقدمة
١٠	ملحق أ (A) : بعض قواعد عمليات الجمع
١٥	ملحق ب (B) : مراجعة للمفاهيم الإحصائية
١٥	متغيرات عشوائية
١٦	دالة احتمال أو كثافة
١٨	الاستقلال وعدم الاستقلال
١٩	نتيجة تمهيدية
١٩	توقعات
٢١	بعض خواص التوقعات
٢٣	عينة عشوائية
٢٤	مقدرات
٢٥	مقدرات غير متحيزة

صفحة

٢٧	اتساق
٢٨	دوال الكثافة المشتركة : إيضاحات
٣١	دوال الكثافة المشتركة : تعميمات
٣٢	دوال الكثافة المشتركة : التوقعات
٣٣	دوال الكثافة المشتركة : توقعات دوال المتغيرات العشوائية
٣٤	توضيح : تغاير X و Y
٣٥	دوال الكثافة المشتركة : مناقشة أكثر عمومية
٣٧	نتيجة مهمة للاستقلال
٣٩	تطبيق شروط الاستقلال

الفصل الثاني : نموذج انحدار المتغيرين

٤١	(١-٢) قياس العلاقة الإحصائية بين متغيرين : التغاير والارتباط
٤٣	التغاير
٤٥	مقدر التغاير
٤٦	عدم تحيز $\hat{\sigma}_{x,y}$
٤٩	اتساق $\hat{\sigma}_{x,y}$
٥٠	تفسير لـ $\hat{\sigma}_{x,y}$
٥٢	معامل الارتباط
٥٨	مقدر معامل الارتباط
٦٠	ملاحظة حول درجات الحرية
٦١	كلمة تحذير
٦١	مثال
٦٥	(٢-٢) وصف العلاقات السلوكية
٧١	(٣-٢) نموذج انحدار المتغيرين
٧٢	الافتراضات الأساسية
٧٧	(٤-٢) تقدير معادلة الانحدار : طريقة المتغير المساعد

مثال	٨٥
ملاحظة على أحد الافتراضات	٨٩
(٥-٢) خواص \hat{a} و \hat{b}	٩١
عدم التحيز	٩٢
تباينات \hat{a} و \hat{b} : بعض الأساسيات	٩٥
تباين المقدرات	٩٨
خاصية أصغر تباين	١٠٢
مقدرات التباين	١٠٣
مثال	١٠٥
خاصية أصغر المربعات لـ \hat{a} و \hat{b}	١٠٦
(٦-٢) قياس القوة التفسيرية لنموذج الانحدار	١٠٨
معامل التحديد	١١٠
$R^2 = \hat{\rho}_{y,\hat{y}}^2$	١١٥
مثال	١١٧
(٧-٢) توضيح : تقدير دالة تكلفة	١١٩
ملحق : إثباتات لثلاث نتائج	١٢١
تباين مجموع المتغيرات العشوائية	١٢١
المقدرات ذات أصغر تباين لـ a و b	١٢٣
خاصية أصغر المربعات لـ \hat{a} و \hat{b}	١٢٦
الفصل الثالث : تطبيقات نموذج الانحدار	
(١-٣) اختبار الفرضيات وفترات الثقة : مقدمة	١٣١
افتراض إضافي	١٣٣
اختبار $b = b_0$ مقابل $b \neq b_0$: مع معرفة σ_u^2	١٣٧
اختبار الفرضيات : تفسير	١٣٩
مناطق القبول والرفض	١٤٠

صفحة

١٤١	فترات الثقة : تفسير
١٤١	بعض الملاحظات على الأخطاء من النوع الأول والنوع الثاني
١٤٤	الفرضية $b \neq 0$
١٤٦	الفرضيات $b < 0$ و $b > 0$
١٤٨	اختبار الفرضيات مع عدم معرفة σ_u
١٤٩	بعض الأمثلة
١٥١	نسبة t : قاعدة للحساب
١٥٤	(٢-٣) مشكلة شكل الدالة
١٥٤	منحنى فليس والتحويل العكسي
١٦٠	التحويل اللوغارتمي
١٦٤	التحويل شبه اللوغارتمي
١٦٨	استخدام التحويلات : تعميمات
١٧٠	(٣-٣) الترجيح ووحدات القياس
١٧٥	مثال
١٧٧	(٤-٣) استخدام المتغيرات المبطة
١٨١	مثال
١٨٣	(٥-٣) التنبؤ
١٨٥	تقدير Y_f^m
١٨٨	التنبؤ بـ Y_f
١٩١	(٦-٣) مثال : التقدير لمنحنى طلب

الفصل الرابع : تحليل الانحدار المتعدد

٢٠٢	(١-٤) نموذج الانحدار المتعدد
٢٠٥	(٢-٤) التقدير بواسطة المتغيرات المساعدة
٢٠٦	المعادلات الطبيعية
٢٠٩	مشكلة تعدد العلاقات الخطية التام

٢١٢	(٣-٤) خصائص المقدرات واختبار الفرضيات
٢١٢	تفسير المقدرات
٢١٥	تباينات المقدرات
٢١٦	فترات الثقة واختبار الفرضيات : بعض المقدمات
٢١٨	فترات الثقة واختبار الفرضيات
٢١٩	(٤-٤) معامل التحديد المتعدد
٢٢٠	R^2 لحالة الانحدار المتعدد
٢٢٢	تعليق على R^2
٢٢٤	معامل التحديد المعدل \bar{R}^2
٢٢٨	(٥-٤) تحليل الانحدار المتعدد - توضيحان
٢٢٨	دالة استهلاك متعددة المتغيرات
٢٣٠	دراسة لضرائب المدينة
٢٣٤	ملحق أ (A) : خصائص المقدرات
٢٣٧	مقدرات غير متحيزة
٢٣٨	تباينات المقدرات
٢٣٩	ملحق ب (B) : العلاقة بين R^2 و \bar{R}^2
	الفصل الخامس : طرق أخرى لتحليل الانحدار المتعدد
٢٤٣	(١-٥) تقدير العلاقات المبطأة
٢٤٧	إبطاء كويك
٢٥٢	إبطاء ألون
٢٦٠	مثال
٢٦٢	(٢-٥) استخدام المتغيرات الصورية
٢٦٨	مثال
٢٧٠	بعض النتائج الإضافية
٢٧٣	(٣-٥) الشكل الدالي مرة أخرى

صفحة

٢٧٣	التحويل اللوغارتمي المعمم
٢٧٧	اشكال متعددة الحدود للمتغيرات المستقلة
٢٨٣	توليفات من الأشكال الدالية
٢٨٥	(٤-٥) توضيح: الطلب على النقود
٢٨٩	ملحق ١ (A): قيود طرفية في ابطاء آلمون
٢٩٢	ملحق ب (B): اختبار الفرضيات المتضمنة لأكثر من معلمة انحدار

الفصل السادس : مشاكل في تحليل الانحدار

٣٠٢	(١-٦) تعدد العلاقات الخطية
٣٠٤	تعدد العلاقات الخطية غير التام : بعض النتائج المنطقية
٣٠٥	تعليق إضافي
٣٠٧	بعض الحلول
٣٠٩	تأثيره على التنبؤ
٣١٠	(٢-٦) مشكلة الارتباط الذاتي
٣١٣	نموذج للانحدار الذاتي
٣١٦	تأثيره على تباينات المقدرات
٣١٧	الوسط الحسابي للمقدرات
٣١٨	طريقة تقدير معممة
٣٢٥	حالة نموذج الانحدار المتعدد
٣٢٦	اختبار درين - واتسون للارتباط الذاتي
٣٣٠	تطبيق
٣٣٥	الارتباط الذاتي والمتغيرات التابعة المبطة
٣٣٨	(٣-٦) اختلاف التباين
٣٣٩	نموذج أساسي
٣٤٢	تأثيره على مقدراتنا
٣٤٣	طريقة للتقدير

صفحة

اختلاف التباين : طرق إضافية للمعالجة	٣٤٦
اختبار لاختلاف التباين	٣٥١
اختبار آخر لاختلاف التباين : اختبار جولد فيلد - كوندات	٣٥٤
بعض التعليقات حول اختبائي اختلاف التباين	٣٥٩
اختلاف التباين : نتيجة للتجميع	٣٦١

(٤-٦) مشاكل في اختيار المتغيرات	٣٦٦
متغير محذوف	٣٦٦
متغيرات أكثر من اللازم	٣٦٩
تعقيبات إضافية	٣٧٠
ملحق : ملاحظة حول الاستقرار	٣٧٣

الفصل السابع : نظم المعادلات

(١-٧) تحيز المعادلات الآنية	٣٧٩
(٢-٧) طريقة المربعات الصغرى ذات المرحلتين : حالة مبسطة	٣٨٤
توضيح : المقدرات المتسقة	٣٨٥
بعض النتائج الإضافية	٣٩٠
(٣-٧) نظم المعادلات : مناقشة أكثر عمومية	٣٩٣
تحديد النموذج	٣٩٣
طبيعة المتغيرات المحددة مسبقاً	٣٩٦
المعادلات الهيكلية ومعادلات الشكل المختزل	٣٩٨
(٤-٧) طريقة المربعات الصغرى ذات المرحلتين : تعميم	٤٠٠
نظرة عامة	٤٠٠
تأطير	٤٠١
م ص م والمتغيرات المحذوفة	٤٠٤
(٥-٧) مشكلة التمييز	٤٠٨

صفحة

٤٠٨	مثال (١)
٤١٢	مثال (٢)
٤١٣	مثال (٣)
٤١٥	عرض أكثر عمومية
٤١٩	بيان عام
٤٢٠	(٦-٧) تقدير م ص م : مثالان
٤٢٠	نموذج للطلب والعرض
٤٢٤	نموذج للمالية العامة المحلية
٤٣١	ملحق : الأخطاء العشوائية المرتبطة ذاتيًا في نموذج المعادلات الآنية

الفصل الثامن : نماذج المعادلات الآنية غير الخطية

٤٤٣	(١-٨) الإطار التحليلي
٤٤٣	توضيح : نموذج من معادلتين
٤٤٥	بعض التوضيحات
٤٤٦	توضيح آخر
٤٤٨	تعميم
٤٤٩	(٢-٨) مشكلة التمييز
٤٤٩	توضيح
٤٥٤	تنقيح
٤٥٧	قاعدة لتمييز النماذج غير الخطية
٤٥٩	تبرير القاعدة
٤٦١	تعميم لتبرير قاعدة التمييز
٤٦٤	(٣-٨) تقدير م ص م
٤٦٤	الخطوط العريضة للطريقة
٤٦٨	تبرير لبعض الملاحظات المهمة
٤٧٣	(٤-٨) تباينات العينة الكبيرة

٤٧٤	النموذج
٤٧٦	تحليل النموذج
٤٨١	ملحق: تقدير النماذج غير الخطية في كل من المتغيرات الداخلية والمعلومات
٤٨١	إطار التحليل
٤٨٤	نتيجة أولية
٤٨٦	طريقة التقدير
٤٨٨	اختيار المتغيرات المستقلة للمرحلة الأولى
٤٨٩	استعراض عام وخطوط عريضة للطريقة
٤٩٠	اختبار الفرضيات وفترات الثقة وتباينات العينة الكبيرة: تعليق
٤٩١	الجدول الإحصائية
٤٩٩	إجابة الأسئلة
	ثبت المصطلحات العلمية
٥٢٩	أولاً: عربي / إنجليزي
٥٣٧	ثانياً: إنجليزي / عربي
٥٤٥	كشاف الموضوعات

مقدمة

إن أحد الأنشطة الأساسية لأي علم هو إختبار المنظم للنظرية في مواجهة الواقع. وعلم الاقتصاد ليس استثناء من هذه القاعدة. فضلا عن ذلك فإن من أكثر التطورات في الاقتصاد في الحقبة الحديثة هو التأكيد المتزايد على تطوير الطرق الإحصائية واستخدامها في تحليل المشكلات الاقتصادية. ويعبر، عادة، عن تلك العلاقات النظرية بين المتغيرات الاقتصادية في شكل رياضي، ولكن لإعطاء هذه العلاقات مضمونا عمليا فقد تزايد استخدام الاقتصاديين لطرق التحليل الإحصائي بهدف اختبار الفرضيات الخاصة بهذه العلاقات، وتقدير أحجامها الفعلية واستخدام هذه التقديرات لعمل تنبؤات كمية للظواهر الاقتصادية. هذا النوع من التحليل هو ما يسمى بالاقتصاد القياسي.

في خطاب وداعي عام ١٩٣٧م وبمناسبة انتهاء عمله كمدير لمدرسة لندن للاقتصاد أعلن اللورد وليام بيفردج William Beveridge منتقدا مهنة الاقتصاد أنه «لفترة مائة سنة في الاقتصاد القياسي تم التعامل مع الحقائق ليس لاختبار النظرية وإنما لتوضيحها... لا يمكن أن يوجد علم إجتماعي حتى تصبح الحقائق المتعلقة بالمجتمع متاحة». ومنذ الفترة التي تلت عبارة بيفردج، حدثت تطورات مهمة في مجالي تطوير الطرق الكمية للتحليل، وتراكم البيانات التي يمكن عن طريقها اختبار النظريات الاقتصادية. ويصعب على قارئ الدوريات الاقتصادية الآن أن يجد عددا واحدا منها يخلو من المقالات التي يدعم فيها مؤلفوها مناقشتهم بالتحليلات القياسية.

ملاحظة عامة داخل الكتاب

ترد الرموز الواردة في المعادلات مائلة بنط أبيض أو أسود وترد داخل المتن إما مائلة أو عادية بالبنط الأبيض أو الأسود (ذلك حسب مايتاح للتغيير في الأصل المرسل على الدسك في قسم الصف).

يعني هذا، أنه، للحصول على المقدرة على فهم البحوث المعاصرة في الاقتصاد وتقويمها (إضافة إلى المقدرة على عمل البحوث ذاتها)، يصبح من الضروري التعرف على علم الاقتصاد القياسي. وعلى سبيل المثال، فإن النقاش المثير والمهم الذي يدور بين ما يسمى بالنقديين والكيينزيين الجدد حول الفاعلية النسبية لكل من السياسة النقدية والسياسة المالية في التأثير على المستوى الكلي للإنتاج والعمالة هو، أساساً، خلاف حول الحقيقة التالية: ماهية هيكل الاقتصاد ومدى استجابته لهذين النوعين من السياسات. وهذا في حد ذاته موضوع ينبغي أن يحسم على أساس الدلائل العملية، وقد اعتمد المشتركون في هذا الجدل بشدة على استخدام الطرق القياسية في التحليل. وما نريد أن نؤكد هنا هو أنه إذا أراد أحد أن يتبع هذا الجدل وأن يختبر الأدلة المقدمة فعليه أن يتحصل على بعض المعرفة في الاقتصاد القياسي (بما فيها معرفة حدوده واحتمال إساءة استعماله، إضافة إلى تفسير النتائج المترتبة على التطبيق الصحيح له في المشاكل الاقتصادية). إذا، الاقتصاد القياسي هو ذلك الفرع من الاقتصاد الذي يعالج السلوك الاقتصادي باستخدام التحليل الكمي. ولذا، فقد أصبح يخدم وظيفتين حيويتين: الأولى أنه يزودنا بطرق للتحقق من النظريات الاقتصادية أو رفضها. فالنظرية الاقتصادية (أو النموذج في اصطلاح الاقتصاديين) هي مجموعة من التعريفات والافتراضات التي يمكن أن يستخدمها الاقتصادي لتوضيح أنواع معينة من الوقائع events. وتصف النظرية الاقتصادية التي يعبر عنها، عادة، في شكل مجموعة من المعادلات، الآلية التي تتفاعل بها المتغيرات الاقتصادية. فعلى سبيل المثال تنص نظرية سلوك المستهلك على أن الكمية التي سيشتريها المستهلكون من سلعة معينة تعتمد على تفضيلاتهم، دخولهم، سعر السلعة ذاتها وأسعار السلع والخدمات الأخرى. وترشدنا هذه النظرية إلى توقع أنه إذا ارتفع سعر السلعة فستتخفّض، عادة، الكمية المشتراة منها.* وفي الاقتصاد الكلي تجد نظريات تتضمن أن المستوى

* يجب ذكر كلمة «عادة» لأنه من المتخيل، في ظل ظروف معينة، أن يكون أثر الدخل الموجب أقوى من أثر الإحلال السالب الناتج عن ارتفاع السعر، ولذا، فإن المستهلك قد يزيد في الواقع مشترياته من السلعة التي ارتفع سعرها.

الإجمالي للاستثمار يعتمد على سعر الفائدة، وبالتحديد تشير هذه النظريات إلى أن معدلات الفائدة الأعلى ستقلل من الإنفاق على تكوين رأس المال الحقيقي (الاستثمار).

ولتقويم فائدة هذه النظريات، يجب أن نحدد مدى الثقة في قدرتها على التنبؤ بالوقائع الاقتصادية. وكما هو الحال في الأمثلة السابق ذكرها، توضع النظرية الاقتصادية في شكل يمكن اختباره عن طريق تحديد ضمني لتتابع سببي من الوقائع مثل: إذا حدث هذا فإن ذلك سيحدث (على سبيل المثال إذا ارتفعت معدلات الفائدة ينخفض الإنفاق الاستثماري). وكثيرا يعبر عن هذا بوساطة المصطلحات الرياضية عن طريق ملاحظة أن متغيرا ما هو دالة في متغير آخر. مثلا نقول إن $I = a - bR$ حيث I ترمز إلى مستوى الإنفاق الاستثماري، R إلى معدل الفائدة، و a ، b هي ثوابت رقمية تأخذ قيما موجبة. في مثل هذه الصيغة يمكن إجراء إختبار التجريبي لصحة توقعات النظرية.

قد نذكر هنا أن اختبار النظريات الاقتصادية بهذه الطريقة ليس عادة عملا سهلا. ذلك أن العبارات السببية من النوع الموصوف أعلاه، عادة ماتبنى على أساس افتراض ثبات العوامل الأخرى على حالها. فالمقولة بأن المعدلات الأعلى من الفائدة تؤدي إلى مستويات استثمار اقل مبنية، مثلا، على افتراض أن الطلب الكلي (من بين أشياء أخرى) يظل ثابتا. فإذا كان الطلب متزايدا في الوقت الذي يتزايد فيه معدل الفائدة فإن ارتفاعا في الاستثمار (لإشباع الطلب المتزايد) قد يكون مصاحبا لمعدلات الفائدة الأعلى. ولا يعني هذا بالضرورة رفض النظرية لأن الأثر السالب للزيادة في معدلات الفائدة قد يقابل بتأثير موجب أعلى للطلب الإجمالي. والمشكلة التي يواجهها الاقتصاديون في هذا المجال هي أن معظم بياناتهم وإحصاءاتهم تأتي من الخبرة اليومية وليست من تجارب معملية متحكم فيها. لهذا السبب فإن على الاقتصاديين القياسيين ابتكار الطرق الإحصائية التي يمكنهم بوساطتها اصطناعيا إبقاء الآثار الأخرى (على المتغير موضع الاهتمام) ثابتة. وبهذه الطريقة، يمكنهم تحديد تأثير متغير على آخر، وهذه المشكلة - وكما

سيوضح في الفصول التالية - هي التي تساعد في اضافة طبيعة خاصة على الطرق الكمية في الاقتصاد.

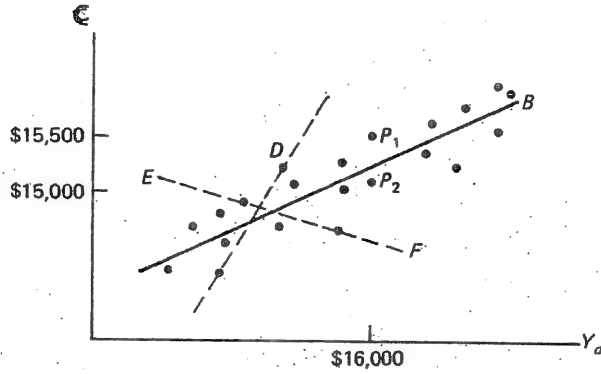
للاقتصاد القياسي، إذن، أهمية أساسية في التحقق من النظريات الاقتصادية أو رفضها. أما الوظيفة الحيوية الثانية للاقتصاد القياسي فهي تزويدنا بتقديرات كمية لأحجام العلاقات بين المتغيرات. فقد تقترح النظرية أن ارتفاعا في السعر يؤدي إلى انخفاض في الكمية المطلوبة. أو أن انخفاضا في مستويات الضرائب يحفز كل من الإنفاق الإجمالي والإنتاج الإجمالي. وعلى الرغم من أن معرفة الطبيعة العامة لهذه العلاقات قيمة جدا، إلا أنها لا تكون مناسبة جدا لأغراض اتخاذ القرارات الفعلية. كما يحتاج رجل الأعمال لمعرفة مقدار النقص في مبيعاته إذا رفع السعر بنسبة ١٠٪. مثلا حتى يتمكن من تقدير تأثير هذا القرار على مستوى ارباحه. وبالمثل فإن المستشار الاقتصادي يجب أن يقدر حجم الزيادة المتوقعة في الإنفاق الإجمالي نتيجة انخفاض محدد في الضرائب. فإذا كان التخفيض الضريبي صغيرا جدا فإنه قد لا يمكن التخفيف من معدلات البطالة المرتفعة ومن الطاقات الإنتاجية المعطلة. بينما إذا كان التخفيض الضريبي كبيرا جدا فقد ينجم عنه التضخم. لهذه الأسباب، يجب أن تكون الطرق الكمية في الاقتصاد قادرة على توليد تقديرات لحجم هذه العلاقات إضافة إلى تقدير العلاقات الأكثر عمومية التي تقترحها النظريات الاقتصادية.

ومن المفيد النظر إلى مثال محدد عند عرض الطبيعة العامة للمشكلة القياسية، وبالمناسبة فإن هذا المثال مهم جدا. افترض - كما اقترح من قبل - أننا مستشارون اقتصاديون قد أوكل إلينا مهمة تقدير حجم الزيادة في الإنفاق الاستهلاكي الناجمة عن إنخفاض محدد مقترح في ضرائب الدخل الشخصية. وكنقطة انطلاق لمهمتنا، يمكننا أن نختار دالة الاستهلاك الكينزية الشهيرة مرجعا نظريا لدراستنا، والتي تقرر أن الإنفاق الاستهلاكي يعتمد على مستوى الدخل المتاح. ويفترض من أجل التبسيط - في الأقل، لتقديراتنا الأولية - أن هذه العلاقة تأخذ الشكل الخطي:

$$C=a+bY_d \quad (1.1)$$

حيث ترمز C إلى الإنفاق الاستهلاكي و Y_d إلى الدخل المتاح و a, b معلمات (ثوابت رقمية). من الواضح أن قيمة المعلمة b لها أهمية كبيرة لنا. حيث إن التخفيض الضريبي سيزيد الدخل المتاح الذي سيحفز بدوره الإنفاق الاستهلاكي و b كما هو معروف هي الميل الحدي للاستهلاك (م ح س)، التي تشير إلى ذلك الجزء من الدولار الإضافي من الدخل المتاح الذي سيوجهه الفرد إلى الاستهلاك. ومن الواضح أننا نحتاج إلى تقدير b إذا رغبنا في تقويم تأثير تخفيض الضرائب على مستوى الإنفاق.

تمدنا النظرية الاقتصادية الكلية ببعض الخطوط العريضة المرشدة والتقريبية. فتقترح النظرية - على سبيل المثال - أن قيمة الميل الحدي للاستهلاك b يتراوح بين الصفر والواحد الصحيح، وأن الدولار الإضافي في الدخل المتاح سيؤدي إلى بعض الزيادة في الإنفاق الاستهلاكي، ولكن جزءاً من هذه الزيادة في الدخل المتاح سوف يدخر، أيضاً، ولذلك فإن الزيادة في الاستهلاك ستكون أقل بعض الشيء من الزيادة في الدخل المتاح. لكننا نحتاج بالطبع إلى تقدير أفضل من ذلك، لأن تأثير تخفيض الضرائب على الإنفاق الاستهلاكي سيكون أكبر إذا كانت $b = 0.9$ (أي إذا كان المستهلكون سينفقون ٩٠٪ من الدخل الإضافي على الاستهلاك ويدخرون فقط ١٠٪) عما لو كانت $b = 0.5$. وهكذا تصبح مهمتنا الأولية هي تحديد قيمة مقدرة لـ b . إحدى الطرق الممكنة للقيام بهذا التقدير هي اختبار السلوك الاستهلاكي والادخاري لمجموعة من الأفراد ذوي المستويات المختلفة من الدخل المتاح. وعلى افتراض حصولنا على هذه المعلومات حول الإنفاق الاستهلاكي والدخل المتاح من خلال استبيان حول ميزانيات الأسرة، وأن هذه المعلومات وضعت في جدول مبين في الشكل (١-١) - يطلق على مثل هذا الشكل والذي سنستخدمه بصورة متكررة شكل الانتشار scatter diagram حيث تمثل كل نقطة فيه قيمتين مشاهدين للمتغيرين. ففي الشكل رقم (١-١) - على سبيل المثال - تشير نقطة P_1 إلى أسرة (في إلاستيان) لها دخل متاح قدره ١٦٠٠٠ دولار وإنفاق استهلاكي قدره ١٥٥٠٠ دولار.



شكل رقم (١-١): بيانات من ميزانيات مفترضة

لنحلل الآن المعلومات المعطاة في الشكل رقم (١-١) في ضوء دالة الاستهلاك الموجودة في معادلة (1.1). نلاحظ أولاً أن التحديد الرياضي لدالة الاستهلاك مؤكد exact، حيث تبين المعادلة (1.1) وجود مستوى محدد من الإنفاق الاستهلاكي يصاحب كل مستوى من مستويات الدخل المتاح. ولكن السلوك الإنساني، بالطبع، يتعد، تماماً، عن مثل هذه الدقة. وفي الحقيقة تبين النتائج الموجودة في الشكل (١-١) أن الأسر ذات مستويات الدخل المتساوية تقوم في معظم الأحوال بانفقات استهلاكية مختلفة. فنقاط مثل P_1 و P_2 - على سبيل المثال - تشير إلى أسرتين كل منهما تحقق مستوى الدخل نفسه وهو ١٦٠٠٠ دولار، ولكن الأسرة الأولى توجه ١٥٥٠٠ دولار منها إلى الإنفاق الاستهلاكي بينما تنفق الأسرة الثانية مبلغ ١٥٠٠٠ دولار فقط، وتدخر ١٠٠٠ دولار.

والسؤال الآن هو كيف يمكننا تحديد تقديرات a و b من هذا الكم الهائل غير المتناسق ظاهرياً من المعلومات. نعرف، من مبادئ الرياضيات، أن نقطتين تحددان الخط المستقيم. ولذلك، فإن المعلومات الضرورية اللازمة لتحديد a و b في المعادلة

(1.1) تتمثل في مشاهدين، فقط. لذا، فمن ناحية، تبدو المشكلة التي تواجهها هي وجود كم كبير من المعلومات، ومن ناحية أخرى، فإن من غير المنطقي (بالفعل) أن نهمل المعلومات الملائمة، وبالطبع، نحصل على قيم مختلفة لكل من a و b إذا اختلفت النقطتان المستخدمتان للوصول إلى حل لهاتين المعلمتين. ففي الشكل رقم (١-١) على سبيل المثال نحصل على الخطين CD و EF أو أي مجموعة أخرى من الخطوط، بحسب النقطتين المختارتين لتحديد الخط.

لكن الفحص الدقيق لانتشار النقاط في الشكل رقم (١-١) يشير إلى وجود نوع من العلاقة بين C و Y_h وهي أنه كلما زاد الدخل المتاح يبدو، أيضا، أن مستوى الإنفاق الاستهلاكي يزداد في المتوسط. وهذه العلاقة بين C و Y_h ليست مؤكدة. غير أنه، على الرغم من ذلك، توجد علاقة نمطية $typical$ بين المتغيرين. يمكننا، ببساطة عن طريق الفحص أو عن طريق منهج أكثر تقدما، أن نوفق خطا مثل AB لنقاط الانتشار هذه. ومثل هذا الخط يعبر عن السلوك الإنفاقي النمطي الذي يمدنا بتقديرات لـ a و b اللتين تشيران إلى القاطع الرأسي وميل الخط المستقيم على التوالي.

تلك هي نوعية المشاكل التي يهتم بها الاقتصاد القياسي. وفي الحقيقة، سيكون تركيز الفصول التالية على تطوير طرق منتظمة ومعقولة لتقدير هذه العلاقات المعتادة بين متغيرين (وفيما بعد بين عدة متغيرات). أثناء مناقشتنا هذه المواضيع، سنكتشف أن هناك عددا من الأسئلة ترتبط بهذه العلاقات ينبغي الإجابة عليها. ففي الحالة الافتراضية السابقة (التي طلبنا أن يتخيل القارئ نفسه في دور المستشار الاقتصادي للتخفيض المقترح في الضرائب)، يتضح، في الحال، وجود مشاكل أخرى عديدة ينبغي حلها إضافة إلى تقدير b قبل الوصول إلى تنبؤ يمكن الإعتماد عليه لتأثير تخفيض الضرائب على الإنفاق الإجمالي. ولإكمال مقدمتنا، فقد يكون من المفيد استعراض لبعض هذه المشاكل ومناقشتها بإيجاز طالما أن حل تلك المشاكل هو مهمة هذا الكتاب.

١ - اختبار الفرضية

افترض أن لدينا نظرية تتضمن علاقة سببية بين متغيرين، بفرض وجود بيانات حول هذين المتغيرين وربما متغيرات أخرى ذات علاقة أيضا. يصبح السؤال هو كيف يمكننا، بدرجة معينة من الثقة، تقرير وجود علاقة بين هذه المتغيرات؟ فعلى سبيل المثال، وبدلالة شكل الانتشار (١-١)، يمكن أن نسأل إلى أي مدى يمكن أن تكون العلاقة الظاهرة بين C و Y_d علاقة زائفة *spurious* وأنها، ببساطة، نتيجة غريبة لهذه العينة بالذات.

٢ - تقدير المعلومات

إذا كانت هناك علاقة بين متغيرين، كيف نتمكن من الاستخدام الأمثل للبيانات المتاحة من أجل الحصول على تقديرات دقيقة لتلك العلاقة؟ بالاعتماد على المعلومات الموجودة في الشكل رقم (١-١) ما الطريقة الأكثر فاعلية لتوليد تقديرات لـ a و b إضافة إلى ذلك نرغب في معرفة مقدار اختلاف السلوك الاقتصادي المتوقع عن المتوسط حتى تكون لدينا فكرة عن مدى فائدة تقديراتنا a و b .

٣ - استخدام التقديرات للتنبؤ

تحت أي مجموعة من الشروط أو القيود يمكننا استخدام هذه التقديرات لعمل تنبؤات؟ بالإشارة، مرة أخرى، للشكل رقم (١-١)، ما الافتراضات التي يجب افتراضها لاستخدام القيمة المقدرة لـ b من المسح العام لموازنات الأسر لتقويم تأثيرات تخفيض الضرائب على الإنفاق الاستهلاكي؟ أو (في موضوع مرتبط بذلك) هل يمكننا التنبؤ على أساس هذه المعلومات؟ وبدرجة ثقة معينة، كم سيكون حجم C عندما يكون مستوى Y_d محددا عند مستوى معين؟

٤ - الصيغة الدالية

ما الشكل الدالي الملائم لهذه العلاقة؟ افترضنا، للتبسيط، وجود علاقة

خطية بسيطة بين C و Y_d ، ولكن، بالتأكيد، ليس من الضروري أن يكون ذلك صحيحا. فرمما يتناقص الميل، الحدي للاستهلاك (أي b) في المعادلة (1.1) مع زيادة الدخل، حيثئذ قد تكون العلاقة الصحيحة هي $C = a + bY_d^{1/2}$. كيف نستخدم النتائج النظرية والبيانات المتاحة لاختبار الشكل الدالي للمتغيرات التي نقدرها؟

٥ - قصور البيانات

ما تأثير القصور في البيانات المتاحة (مثلا، أخطاء القياس) على نتائجنا؟ هل تؤدي إلى عدم صحة تقديراتنا؟

٦ - علاقات التغذية المرتدة

افترض أننا نرغب في تقدير تأثير المتغير X على المتغير Y . ففي بعض الأحيان من الممكن أن X لا يؤثر، فقط، على Y ولكن Y أيضا يؤثر في X . في هذه الحال، قد يكون من الصعب التمييز بين: هل تعكس تقديرات المعلمة تأثير X على Y أو ربما على الأرجح أن تكون مزيجا من هذين التأثيرين معا. وكثيرا ماتحدث علاقات التغذية المرتدة هذه حدوثا متكررا في الاقتصاد، فمثلا يحدد السعر الكمية المطلوبة من السلعة والكمية المطلوبة تؤثر بدورها في السعر، وكذلك فإن لمستوى الإنفاق الإجمالي في الاقتصاد تأثيرا قويا على مستوى الناتج الكلي والدخل الكلي ولكن مستوى الناتج والدخل يؤثر، بدوره، في مستوى الإنفاق... وهلم جرا.

وهذه الظاهرة هي ماقد يطلق عليها اسم مشكلة النظم. والمشاكل الاقتصادية غالبا ماتكون من نوع مشكلة النظم، وتعكس الاعتماد المتبادل الذي يميز عادة، عمل النظام الاقتصادي. ولكن هذه، كما سنرى، توجد مشاكل خطيرة للاقتصاد القياسي الذي ينبغي عليه أن يحاول فض اشتباك هذا الاعتماد المتبادل كليا.

هذه هي بعض مشاكل الاقتصاد القياسي، وسنطور في الفصول التالية طرق معالجتها.

ملحق أ (A): بعض قواعد عمليات الجمع

كما أشرنا في المقدمة، لا يستخدم هذا الكتاب نظريات متقدمة في الرياضيات أو الإحصاء، ولذا، فسوف نعتمد كلية في التحليل على المبادئ الأولية للجبر والإحصاء. وعلى الرغم من ذلك، يوجد عدد قليل من القواعد المرتبطة بالعمليات الجبرية التي تبدو إما جديدة أو مبهمة لبعض القراء. وبما كنا سنستخدم تلك القواعد استخداما مكثفا فقد يكون من الأسهل للتحليل أن نعرض لها هنا حتى يتعود القارئ عليها من البداية.

سوف نستخدم سيجما (Σ الكبيرة) لترمز إلى عملية الجمع، على سبيل المثال إذا رمزنا إلى الكمية المنتجة من إحدى السلع في السنة الأولى بالرمز Q_1 أو بعمومية أكثر إذا جعلنا Q_t ترمز إلى الكمية المنتجة من السلعة في السنة t ، حيث t فإن إجمالي الكمية المنتجة في السنوات الأولى والثانية والثالثة يمكن أن نعبر عنها $(Q_1 + Q_2 + Q_3)$. ويمكن كتابة هذا المقدار باختصار، على النحو التالي:

$$\sum_{t=1}^3 Q_t \text{ حيث:}$$

$$\sum_{t=1}^3 Q_t = Q_1 + Q_2 + Q_3 \quad (1A.1)$$

ويمكننا أن نعمم النتيجة السابقة بإستخدام المقدار $\sum_{t=1}^n Q_t$ ليعبر عن مجموع

الحدود الأولى التي عددها n من المتغير Q . وعلى سبيل المثال، لتوضيح هذه الفكرة فإنه يمكننا أن نعبر عن مجموع الحدود الجبرية من الحد الثالث إلى الحد السابع من المتغير Q على النحو التالي:

$$\sum_{t=3}^7 Q_t = Q_3 + Q_4 + Q_5 + Q_6 + Q_7 \quad (1A.2)$$

وقبيل الاستمرار، عليك أن تثبت مايلي:

$$\sum_{t=1}^{16} Q_t - \sum_{t=3}^{17} Q_t = Q_1 + Q_2 - Q_{17}. \quad (1A.3)$$

والآن سوف نستعرض بعض قواعد الجمع التي تستخدم استخداما متكررا في هذا الكتاب.

القاعدة الأولى

إذا كانت c مقداراً ثابتاً (مثل $c = 5$) فإن:

$$\sum_{t=1}^n cX_t = c \sum_{t=1}^n X_t.$$

ولتوضيح ذلك نعرف، أن $\sum_{t=1}^n cX_t$ هي مجموع أعداد قدرها n من القيم الأولى للمتغير X ، التي يكون كل منها مضروباً في مقدار ثابت قدره c ، ولذا فإن:

$$\sum_{t=1}^n cX_t = cX_1 + cX_2 + \dots + cX_n = c(X_1 + X_2 + \dots + X_n) = c \sum_{t=1}^n X_t. \quad (1A.4)$$

القاعدة الثانية

إذا كانت كل من X و Y متغيرات Variables، فإن:

$$\sum_{t=1}^n (X_t + Y_t) = \sum_{t=1}^n X_t + \sum_{t=1}^n Y_t$$

وتعني هذه القاعدة أن مجموع قيم X و Y هو مجموع قيم X مضافاً إليها مجموع قيم Y ، وإثبات ذلك على النحو التالي:

$$\begin{aligned}
\sum_{t=1}^n (X_t + Y_t) &= (X_1 + Y_1) + (X_2 + Y_2) + \dots + (X_n + Y_n) \\
&= (X_1 + X_2 + \dots + X_n) + (Y_1 + Y_2 + \dots + Y_n) \\
&= \sum_{t=1}^n X_t + \sum_{t=1}^n Y_t
\end{aligned} \tag{1A.5}$$

لتعميم القاعدتين الأولى والثانية فإن عليك أن تثبت:

$$\sum_{t=1}^n (aX_t + bY_t + cZ_t) = a \sum_{t=1}^n X_t + b \sum_{t=1}^n Y_t + c \sum_{t=1}^n Z_t,$$

حيث a, b, c ثوابت و X, Y, Z متغيرات.

القاعدة الثالثة

إذا كانت \bar{X} هي المتوسط الحسابي لعدد n من قيم المتغير X ، حيث

$$\bar{X} = (\sum_{t=1}^n X_t) / n \quad \text{فإن:}$$

مجموع الانحرافات العن
عن وسط الحسابي
يساوي الصفر

$$\sum_{t=1}^n (X_t - \bar{X}) = 0$$

ولإثبات هذه القاعدة، عليك أن تلاحظ أولاً:

$$\sum_{t=1}^n (X_t - \bar{X}) = \sum_{t=1}^n X_t - \sum_{t=1}^n \bar{X}. \tag{1A.6}$$

فإذا ضربنا وقسمنا الحد الأول من الطرف الأيمن على n فسنحصل على:

$$\frac{\sum_{t=1}^n X_t}{n} = n\bar{X} \tag{1A.7}$$

بعد ذلك، نلاحظ أن:

$$\sum_{t=1}^n \bar{X} = \bar{X} + \bar{X} + \dots + \bar{X} = n\bar{X}. \quad (1A.8)$$

وبالتعويض من (1A.7) و (1A.8) في (1A.6)، فإننا نحصل على النتيجة المطلوبة:

$$\sum_{t=1}^n (X_t - \bar{X}) = \sum_{t=1}^n X_t - \sum_{t=1}^n \bar{X} = n\bar{X} - n\bar{X} = 0$$

من هذه المناقشة يتضح لنا أنه إذا كانت K مقداراً ثابتاً، فإن:

$$\sum_{t=1}^n K = nK. \quad (1A.9)$$

القاعدة الرابعة

إذا كان كل من \bar{X} و \bar{Y} هو المتوسطات الحسابية لعدد n من القيم للمتغيرين X و Y ، فإن:

$$\sum_{t=1}^n (X_t - \bar{X})(Y_t - \bar{Y}) = \sum_{t=1}^n (X_t - \bar{X})Y_t.$$

ولإثبات ذلك، علينا أن نلاحظ أولاً أن:

$$\begin{aligned} \sum_{t=1}^n (X_t - \bar{X})(Y_t - \bar{Y}) &= \sum_{t=1}^n [(X_t - \bar{X})Y_t - (X_t - \bar{X})\bar{Y}] \\ &= \sum_{t=1}^n (X_t - \bar{X})Y_t - \sum_{t=1}^n (X_t - \bar{X})\bar{Y} \end{aligned} \quad (1A.10)$$

وسوف نوضح الآن أن الحد الثاني من المقدار الجبري في الطرف الأيمن يعادل الصفر:

$$\sum_{t=1}^n (X_t - \bar{X})\bar{Y} = \bar{Y} \sum_{t=1}^n (X_t - \bar{X}) = \bar{Y} \cdot 0 = 0; \quad (1A.11)$$

ويستج ذلك عن القاعدتين الأولى والثالثة (مع ملاحظة أن \bar{Y} ثابت). وهذا هو المطلوب لإثبات القاعدة الرابعة.

ويمكن أن نستمرسل في تحليل الخطوة السابقة من خلال ملاحظة أن:

$$\sum_{i=1}^n (X_i - \bar{X})Y_i = \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n \bar{X} Y_i. \quad (1A.12)$$

فإذا ما ضربنا وقسمنا الحد الأخير بـ n فإنه يمكننا أن نعبر عن المعادلة (1A.12)

على النحو التالي:

$$\sum_{i=1}^n (X_i - \bar{X})Y_i = \sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y}. \quad (1A.13)$$

ونترك للقارئ أن يثبت هاتين النتيجةين التابعتين Corollaries للقاعدة الرابعة:

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n (X_i - \bar{X})X_i$$

و

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2.$$

تلميح للحل: عبر عن $\sum_{i=1}^n (X_i - \bar{X})^2$ على أساس أنه يعادل المقدار

$$\left(\sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X}) \right)$$

ملحق ب (B) مراجعة للمفاهيم الإحصائية

نعرض في هذا الملحق مراجعة مختصرة لبعض المفاهيم الأساسية في الإحصاء التي سوف تستخدم استخداماً متكرراً في هذا الكتاب. وبالطبع، فإن هذا الملحق لا يقصد به أن يكون بديلاً عن الحصول على دراسة أولية للإحصاء، وإنما يهدف إلى تزويد القارئ بمراجعة عامة ودقيقة في الوقت نفسه لبعض المفاهيم الإحصائية المختارة.

متغيرات عشوائية Random variables

لأغراض عملية، يمكن النظر للمتغير العشوائي على أنه متغير تتحدد قيمته على أساس نتيجة تجربة، بشرط أن تكون النتيجة عرضة للمصادفة. وبمعنى آخر، ترتبط كل قيمة ممكنة للمتغير العشوائي باحتمال معين للحدوث. على سبيل المثال، فإن قيمة متغير عشوائي يمكن أن تعتمد على رمي قطعة عملة معدنية في الهواء ومشاهدة الوجه الذي يظهر منها بعد استقرارها على سطح مستو. وتكون نتيجة إلقاء هذه القطعة من العملة المعدنية (أو إجراء التجربة) إما كتابة (H) أو شعار (T). في هذه الحالة، يمكننا أن نعرف المتغير العشوائي Y على أساس أنه المتغير الذي تكون قيمته مساوية للواحد إذا كان الوجه المشاهد للعملة هو H ، وتكون قيمته مساوية للصفر إذا كان الوجه المشاهد هو T .

ويمكن التعبير عن العبارة السابقة بدقة أكثر واختصار من خلال التصريح بأن Y هو متغير عشوائي يأخذ القيم $y = 0, 1$. وعلينا هنا ألا نخلط بين كل من Y و y ، فالأول Y هو المتغير العشوائي الذي تعتمد قيمته على نتيجة التجربة، بينما y هو إحدى القيم المحددة (أرقام) التي قد يأخذها Y .

وأحد المتغيرات العشوائية الأخرى هو W ، حيث W هو الوزن بالأرطال للشخص الذي اختير عشوائياً من عينة معطاة من الأفراد. في هذه الحالة تكون القيم الممكنة لـ W هي w حيث $50 \leq w \leq 1000$ إذا كانت المجموعة من الأفراد

تتألف من الأفراد البالغين. وعلى الرغم من أن كل من Y و W هي متغيرات عشوائية فإن هناك اختلافا مهما بينهما: حيث يمكن أن تأخذ W أي قيمة في مدى القيم المتصلة Continuous، بينما لا تأخذ قيم Y مثل هذا الاتصال (أو إلستمرارية). ويطلق على المتغيرات العشوائية من النوع W المتغيرات العشوائية المتصلة (أو المستمرة)، بينما يطلق على المجموعة الثانية من المتغيرات العشوائية (من النوع Y) بالمتغيرات العشوائية المتقطعة discrete.

وفي هذا الملحق الإحصائي يأخذ التحليل شكل المتغيرات العشوائية المتقطعة، ويرجع السبب في ذلك إلى أن تحليل المتغيرات العشوائية المتصلة يتطلب استخدام التفاضل والتكامل. ولما كان هذا الكتاب يعتمد في التحليل على المبادئ الأولية في الجبر والإحصاء (كما ذكرنا في الملحق أ (A)) فيكفي لنا لهذا الغرض مناقشة مانحتاجه من المفاهيم المرتبطة بالمتغيرات العشوائية المتقطعة فقط.*

دالة احتمال أو كثافة** Probability (Or Density) function

ترتبط بالمتغير العشوائي دالة احتمال (يطلق عليها أحيانا دالة الكثافة الاحتمالية) وتعطي هذه الدالة الاحتمالات التي يأخذ فيها المتغير العشوائي كل قيمة من القيم الممكنة له. ويعبر، عادة، عن دالة الاحتمال على شكل معادلة أو جدول. ومن مثال القاء قطعة معدنية في الهواء المعطى سابقا، فإن القيم الممكنة للمتغير العشوائي Y هي $y = 0,1$ ، والاحتمالات المرتبطة بها (على افتراض توازن القطعة المعدنية وأن عملية الإلقاء غير متحيزة لأي من الوجهين) هي $1/2$ و $1/2$. لذا، يمكننا كتابة الدالة الاحتمالية لـ Y على النحو التالي: $y = 0,1$ و $f(y) = 1/2$ أي أنه إذا كان Y متغيرا

* القراء الذين تتوافر لديهم خلفية رياضية ملائمة يمكنهم، عموما ترجمة التحليل الموجود في هذا الملحق إلى مجال المتغيرات العشوائية المتصلة عن طريق إحلال علامات التكامل محل علامات الجمع.

** يلاحظ أن المؤلف يستخدم دالة الاحتمال ودالة الكثافة ليدلا على المعنى نفسه، بينما تستخدم في الإحصاء دالة الاحتمال للتعبير عن دوال المتغيرات المتقطعة، وتستخدم دالة الكثافة للتعبير عن دوال المتغيرات المتصلة (ملاحظة المترجم)

عشوائيا لدالة احتمال $f(y)$ فإن العبارة $f(1)$ تعني أن احتمال أن يأخذ المتغير العشوائي Y الواحد الصحيح يكون مساويا لـ $1/2$.

ولمزيد من التوضيح نأخذ مثالا آخر. فإذا افترضنا أن المتغير Z يعبر عن الرقم الذي يظهر على زهر النرد المتزن عند دحرجته. فإن مدى القيم التي يمكن أن يأخذها Z هو: $z = 1, 2, 3, 4, 5, 6$ ، وتكون دالته الاحتمالية على النحو التالي: $g(z) = \frac{1}{6}, z = 1, \dots, 6$. ويمكن التعبير عن هذه المعلومات تعبيرا آخر في شكل الجدول التالي:

Z	1	2	3	4	5	6
g(z)	1/6	1/6	1/6	1/6	1/6	1/6

(1B.1)

مع ملاحظة أن:

$$g(1) + g(2) + \dots + g(6) = \frac{1}{6} + \frac{1}{6} + \dots + \frac{1}{6} = 1 \quad (1B.2)$$

أي أنه لما كان المتغير العشوائي Z يجب أن يأخذ إحدى القيم الصحيحة من $1, 2, \dots, 6$ فإن مجموع الاحتمالات في هذه الحالة يكون مساويا للواحد الصحيح. وباختصار، فإن المعادلة (1B.2) تعني أن احتمال أن يأخذ المتغير Z إحدى القيم من $(1, 2, \dots, 6)$ يكون مساويا للواحد الصحيح.

وتعد هذه النتيجة نتيجة عامة ولا تقتصر، فقط، على هذا المثال. ذلك أن مجموع الاحتمالات المقابلة لكل القيم الممكنة للمتغير العشوائي يجب أن تساوي واحداً صحيحاً. وإحدى النتائج العامة الأخرى هي أنه، لما كان كل احتمال من الاحتمالات يكون إما أكبر من الصفر أو مساويا للصفر، فإن الدالة الاحتمالية يجب أن تعرف، فقط، في حدود ذلك المجال، فعلى سبيل المثال، في حال زهر النرد المشار إليه آنفاً، فإن قيمة الدالة الاحتمالية المناظرة لـ $z = \sqrt{363}$ تساوي الصفر، لأن احتمال $z = \sqrt{363}$ يساوي الصفر.

الاستقلال وعدم الاستقلال Independence and dependence

تنشأ المشاكل الإحصائية، غالبا، من العلاقة التي تربط بين العديد من المتغيرات العشوائية. فعلى سبيل المثال افترض أن زهرا للنرد دحرج مرتين عشوائيا، وأن Z_1 و Z_2 هي القيم التي ظهرت في الدرجة الأولى والدرجة الثانية على الترتيب. ففي هذه الحال لانتوقع أن تكون قيمة Z_2 قد تأثرت بقيمة Z_1 ، أو أن Z_1 قد تأثرت بقيمة Z_2 ، فمثلا، إذا كان زهر النرد متوازنا فإن احتمال الحصول على الرقم 3 في الدرجة الثانية (أو $Z_1=3$) سوف تكون مساوية $1/6$ بغض النظر عن نتيجة الدرجة الأولى. ويمكن أن نعبر عن ذلك بلغة احصائية بالقول إن احتمال أن يأخذ Z_2 أي قيمة لا يتأثر بالقيمة المحددة التي أخذها Z_1 والعكس صحيح. ويقال في هذه الحالة عن المتغيرين Z_1 و Z_2 (والتي تكون احتمالاتها غير مرتبطة بهذه الصورة) أنهما متغيرين عشوائيان مستقلان عن بعضهما بعضا. أما إذا كانت المتغيرات العشوائية تعتمد على بعضها بعضا فيطلق عليها متغيرات عشوائية غير مستقلة. ولتوضيح ذلك، على سبيل المثال، إذا سحبنا ورقة واحدة من أوراق اللعب من مجموعة كاملة من هذه الأوراق وجعلنا $P=1$ إذا كانت الورقة إحدى الصور و $P=0$ إذا كانت الورقة غير ذلك. إضافة إلى ذلك افترض أن $K=1$ إذا كانت الورقة المسحوبة ولدا و $K=0$ إذا كانت غير ذلك. حيثئذ، فإن المتغيرات العشوائية P و K هي متغيرات عشوائية غير مستقلة حيث إن احتمال أن $K=1$ سيكون مساويا $4/12$ إذا كانت $P=1$ ، ويساوي صفرا إذا كانت $P=0$. أما إذا لم تعط أي معلومات مرتبطة بـ P فإن احتمال $K=1$ يعادل $4/52$. وباختصار، نقول إن متغيرين عشوائيين يكونان غير مستقلين إذا كانت المعلومات المرتبطة بأحدهما تغير من الاحتمالات المرتبطة بالآخر.

ومن السهل أن نعمم هذه التعاريف، حيث يكون المتغير العشوائي X_1 مستقل عن المتغيرات العشوائية X_2, \dots, X_n إذا كان احتمال أن يأخذ X_1 أي قيمة لا يتأثر مطلقا بالقيم المحددة التي تأخذها المتغيرات X_2, \dots, X_n ، أما إذا وجد بعض التأثير على X_1 بواسطة القيم التي تأخذها X_2, \dots, X_n ، ففي هذه الحال، يكون X_1 وواحد على الأقل من المتغيرات الأخرى X_2, \dots, X_n غير مستقلة.

نتيجة تمهيدية

نفترض أن X_1, \dots, X_n متغيرات عشوائية وأن Y هي دالة في هذه المتغيرات أي أن:

$$Y = h(X_1, \dots, X_n) \quad (1B.3)$$

تعني الدالة (1B.3) أن Y تعتمد على مجموعة جزئية subset من X 's أو تعتمد على كل المتغيرات X_1, \dots, X_n ، ولما كانت X 's متغيرات عشوائية فإن Y يكون متغيرا عشوائيا أيضا، بمعنى أن قيمته المحددة سوف تعتمد على الصدفة. يضاف إلى ذلك أنه إذا كانت الدالة معقولة reasonable فإن Y سوف تكون لها دالة كثافة احتمالية كما هو الحال لكل من المتغيرات العشوائية X_1, \dots, X_n .

ولأن الشروط المطلوبة لوجود دالة الكثافة الاحتمالية Y والحسابات المتضمنة في تحديدها تتجاوز مستوى هذا الكتاب، فإن النتائج الموجودة في هذا الملحق والمتعلقة بدوال المتغيرات العشوائية (مثل Y) لا تتطلب حساب تلك الدوال وإنما نفترض ضمنا وجودها حتى يمكن افتراض المفاهيم نفسها المتعلقة بالمتغيرات العشوائية X 's للمتغير Y كما سيتضح لاحقا.

توقعات Expectations

يعرف التوقع الرياضي $E(X)$ (وغالبا ما يطلق عليه القيمة المتوقعة) للمتغير العشوائي X ذي القيم الممكنة $x = x_1, x_2, \dots, x_n$ والذي تكون دالته الاحتمالية $f(x)$ ، على النحو التالي:

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + \dots + x_n f(x_n) \quad (1B.4)$$

ومن المعادلة (1B.4)، يمكن تعريف القيمة المتوقعة لـ X بأنها القيمة المتوسطة المرجحة لكل القيم الممكنة للمتغير العشوائي X . حيث الأوزان هي الاحتمالات المناظرة لكل قيمة من قيم X . ويعرف الرمز E في المعادلة السابقة بأنه معامل القيمة المتوقعة the expected value operator، وعلى سبيل مثال لتوضيح كيفية حساب

القيمة المتوقعة نجد أن تلك القيمة للمتغير العشوائي Z في مثال زهرة النرد المشار إليه أنفا تكون مساوية لـ:

$$E(Z) = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + \dots + 6\left(\frac{1}{6}\right) = \frac{21}{6} = 3\frac{1}{2} \quad (1B.5)$$

ويعبر، غالباً، عن القيمة المتوقعة للمتغير العشوائي بالوسط الحسابي mean ويرمز له بالحرف μ ، ويلحق بالحرف μ بوصفه دليلاً سفلياً subscript رمز المتغير العشوائي الذي يتم حساب الوسط الحسابي له، على سبيل المثال فإن: $E(Z) = \mu_Z$ و $E(X) = \mu_X$. ويمكن لنا أن نفكر في الوسط الحسابي للمتغير العشوائي بأنه مقياس نزعة المركزية central tendency أو موقعة. فإذا ما كررت التجربة عدداً من المرات فإن الوسط الحسابي يكون القيمة التي نتوقعها في المتوسط للمتغير على مدى كل التجارب. ويعرف التباين للمتغير العشوائي X (σ_x^2) حيث $x = x_1, x_2, \dots, x_n$ قيمه الممكنة و $f(x)$ دالته الاحتمالية بأنه:

$$\sigma_x^2 = E(X - \mu_x)^2 \quad (1B.6)$$

حيث $E(X) = \mu_x$ ، ومن المعادلة (1B.6)، نرى أن التباين هو القيمة المتوقعة لمربع انحرافات قيم المتغير عن وسطه الحسابي. وبمعنى آخر، فإن التباين هو مقياس لتشتت قيم المتغير العشوائي حول وسطه الحسابي، ويشير في المتوسط إلى مدى بعد قيم المتغير العشوائي عن وسطه الحسابي. وعلى سبيل المثال للحسابات التي يتضمنها إيجاد التباين للمتغير العشوائي، نوجد تباين المتغير Z السابق الإشارة إليه:

$$\begin{aligned} \sigma_z^2 &= E(X - 3.5)^2 \\ &= (1 - 3.5)^2\left(\frac{1}{6}\right) + (2 - 3.5)^2\left(\frac{1}{6}\right) + \dots + (6 - 3.5)^2\left(\frac{1}{6}\right) \\ &= \frac{17.50}{6} = 2\frac{11}{12} \end{aligned}$$

ويعرف الجذر التربيعي الموجب للتباين بالانحراف المعياري the standard deviation.

بعض خواص التوقعات

نعرض في هذا المبحث، باختصار، بعض خصائص التوقعات والتي ستستخدم مرارا في هذا الكتاب. وأولى هذه الخصائص هي أن القيمة المتوقعة للرقم الثابت (c) هي نفسه كمايلي:

$$E(c) = c \quad (1B.7)$$

فإذا كان الرقم الثابت c يعادل خمسة مثلا، فإن المعادلة السابقة تعني أن القيمة المتوقعة لـ c هي خمسة. ويرجع ذلك إلى أنه لما كانت c لا تأخذ قيما أخرى غير ٥ فإنها تأخذ القيمة ٥ باحتمال قدره الواحد الصحيح ولذلك فإن

$$E(5) = 5 \cdot f(5) = 5(1) = 5$$

والآن، افترض أن المتغير العشوائي Y هو حاصل ضرب رقم ثابت في متغير عشوائي آخر، فمثلا في حالة دحرجة زهر النرد افترض أن $Y = 15Z$ ، فإذا اظهر زهر النرد الرقم 4 مثلا فإن المتغير Y يأخذ القيمة $Y = 15(4) = 60$. ويمكننا أن نوجد القيمة المتوقعة Y على النحو التالي:

$$\begin{aligned} E(Y) &= E(15Z) = 15(1)\left(\frac{1}{6}\right) + 15(2)\left(\frac{1}{6}\right) + \dots + 15(6)\left(\frac{1}{6}\right) \\ &= 15(3.5) = 52.5 \end{aligned}$$

وهكذا يتضح لنا أن القيمة المتوقعة لـ $Y = 15Z$ ما هي إلا حاصل ضرب 15 في القيمة المتوقعة للمتغير Z. وهذه نتيجة عامة، فإذا كان لدينا ثابت b ومتغيرا عشوائيا X فإن:

$$E(bX) = bE(X) = b\mu_x \quad (1B.8)$$

ويمكننا أن نتوسع في تطبيق هاتين النتيجةين العامتين على النحو التالي: افترض أن X_1, X_2, \dots, X_n متغيرات عشوائية عددها n وأوساطها الحسابية هي $\mu_1, \mu_2, \dots, \mu_n$ على الترتيب، وأن المتغير Y يعرف على النحو التالي:

$$Y = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n \quad (1B.9)$$

حيث إن a_1, a_2, \dots, a_n ثوابت.

وتعني الدالة (1B.9) أن المتغير هو، في واقع الأمر توليفة خطية Linear combination من X 's فنلاحظ الآن، بدون إثبات، أن:

$$\begin{aligned} E(Y) &= E(a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n) \\ &= E(a_0) + E(a_1X_1) + E(a_2X_2) + \dots + E(a_nX_n) \\ &= a_0 + a_1E(X_1) + a_2E(X_2) + \dots + a_nE(X_n) \\ &= a_0 + a_1\mu_1 + a_2\mu_2 + \dots + a_n\mu_n \end{aligned} \quad (1B.10)$$

فإذا كانت Y دالة خطية في مجموعة من المتغيرات العشوائية فإن القيمة المتوقعة لـ Y هي مجموع القيم المتوقعة للمتغيرات العشوائية التي تتكون منها Y .
افترض الآن أننا أوجدنا متغيراً عشوائياً آخر واطلقنا عليه Q ، حيث إن $Q = Z^2$ ،
وإن Z هي القيمة التي تظهر على زهر النرد عند دحرجته، لذا فإن قيمة Q ماهي إلا مربع العدد الذي يظهر على الزهر. وبأخذ القيمة المتوقعة لـ Q نجد أن:

$$E(Q) = E(Z^2) = 1\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + \dots + 36\left(\frac{1}{6}\right) = 15\frac{1}{6} \quad (1B.11)$$

وبما أن $E(Z) = 3.5$ لذا نرى أن:

$$[E(Z)]^2 = (3.5)^2 = 12.25 \neq E(Z^2) = 15\frac{1}{6}$$

وهكذا فإن $[E(Z)]^2 \neq E(Z^2)$ ولفظياً، فإن مربع القيمة المتوقعة للمتغير العشوائي Z لا يساوي القيمة المتوقعة لمربع المتغير العشوائي Z .

يوضح المثال السابق نتيجة أكثر عمومية وهي: أنه، إذا كان $Y = g(X)$ ، و $g(X)$ دالة غير خطية في المتغير العشوائي X فإن:

$$E(Y) = E[g(X)] \neq g[E(X)] \quad (1B.12)$$

على سبيل المثال، فلقد رأينا توءاً أن $[E(X)]^2 \neq E(X^2)$ ، ومثال ذلك،
أيضاً أن $E(e^x) \neq e^{E(x)}$.

وبخصوص التوقعات، فإننا نحتاج نتيجة أخرى. افترض أن Y تساوي حاصل ضرب مجموعة من المتغيرات العشوائية:

$$Y = (X_1 X_2 \dots X_n) \quad (1B.13)$$

ففي هذه الحال، إذا لم تكن المتغيرات X_1, X_2, \dots, X_n مستقلة عن بعضها بعضاً تكون لدينا النتيجة العامة التالية:

$$E(Y) = E(X_1 X_2 \dots X_n) \neq E(X_1)E(X_2) \dots E(X_n) \quad (1B.14)$$

ولإثبات ذلك، فقد عرفنا من قبل أن:

$$E(Z^2) = E(Z \cdot Z) \neq E(Z)E(Z) = [E(Z)]^2$$

ولكن إذا كانت المتغيرات العشوائية X 's مستقلة عن بعضها بعضاً فإن التوقع الرياضي لحاصل ضرب هذه المتغيرات في بعضها بعضاً يكون مساوياً حاصل ضرب توقعاتها:

$$E(Y) = E(X_1 X_2 \dots X_n) = E(X_1)E(X_2) \dots E(X_n) \quad (1B.15)$$

حيث إن جميع X 's مستقلة.

عينة عشوائية Random sample

لنفرض أن لدينا قطعة عملة معدنية ليست متوازنة تماماً، ففي هذه الحال، يكون احتمال الحصول على الكتابة لهذه القطعة عند رميها هو P ، حيث P ليس بالضرورة مساوياً $1/2$. واحتمال الحصول على الشعار سوف يعادل $(1-P)$. ويطلق على الثابت P الذي يظهر في المعادلة أو في النموذج الاحتمالي معلمة Parameter. والآن، افترض أننا لانعرف قيمة المعلمة P ، ولكننا نريد الحصول على تقدير لها. لعمل ذلك يمكننا أن نلقي بقطعة العملة المعدنية عدداً من المرات (مائة مرة مثلاً) ونأخذ \hat{P} بوصفه تقديراً لـ P حيث \hat{P} هي نسبة ظهور الكتابة إلى عدد الرميات الكلية، أي عدد مرات ظهور الكتابة مقسوماً على مائة. ولتوضيح كيفية عمل ذلك افترض أن X_1 متغير عشوائي يأخذ القيمة صفراً إذا ظهرت الشارة في الرمية الأولى أو يأخذ قيمة الواحد الصحيح إذا ظهرت الكتابة. وافترض أيضاً، أن X_2, \dots, X_{100} هي متغيرات عشوائية تأخذ قيمتها الصفر والواحد الصحيح

على الترتيب وفقا لنتائج الرمي من 2 إلى 100. بمعنى أن $X_i=0$ إذا ظهرت الشارة على الرمية رقم i أو $X_i=1$ إذا أسفرت تلك الرمية عن ظهور الكتابة. وعلى افتراض أن احتمال الحصول على الكتابة في أي رمية ليس مرتبطا بنتيجة أي رمية أخرى، في هذه الحال تكون المتغيرات X_1, X_2, \dots, X_{100} مستقلة. يضاف إلى ذلك أن جميع هذه المتغيرات العشوائية سيكون لها دالة الاحتمال نفسها. بمعنى أن دالة الاحتمال لكل X_i ستكون:

$$\begin{array}{c|c|c} x_i & 0 & 1 \\ \hline f(x_i) & 1-p & p \end{array} \quad (1B.16)$$

وتكون المتغيرات العشوائية المستقلة: مثل X_1, X_2, \dots, X_{100} والتي يكون لها دالة الاحتمال نفسها عينة عشوائية. ويوصف المجتمع الذي سحبت منه هذه العينة العشوائية بدالة الاحتمال المشتركة لتلك المتغيرات العشوائية. وفي هذه الحالة، تكون دالة الاحتمال للمجتمع الإحصائي على النحو التالي:

$$\begin{array}{c|c|c} x & 0 & 1 \\ \hline f(x) & 1-p & p \end{array} \quad (1B.17)$$

مقدرات Estimators

يمكن وصف الطريقة التي قدرنا بها P (في المثال السابق) بإيجاد قيمة \hat{P} بدلالة المتغيرات العشوائية X_1, X_2, \dots, X_{100} على النحو التالي:

$$\hat{P} = \sum_{i=1}^{100} \frac{X_i}{100} \quad (1B.18)$$

فإذا ظهرت الكتابة، على سبيل المثال، ثمانين مرة من مائة الرمية لقطعة العملة فإن 80 من الـ X_i ستكون واحدا بينما تكون قيمة الرميات الـ 20 الأخرى صفرا، ولذا، فإن P ستصبح 80/100 أو 0.8.

وبلغة الإحصاء، يطلق على الرقم 0.8 في المثال أعلاه تقدير estimate المعلمة

P ، بينما يطلق على القاعدة أو الصيغة الرياضية التي تستخدم للحصول على هذا التقدير (\hat{P}) في الحالة أعلاه المقدّر. ويعني ذلك أن التقدير رقم معين محسوب على أساس المقدّر، فعلى سبيل المثال، توضح لنا P أعلاه أننا نضيف قيم المتغيرات العشوائية المائة X_1, X_2, \dots, X_{100} ومن ثم تقسمها على 100 وهذا هو مقدّرنا، ولكن، للحصول على تقدير معين لـ P (مثلاً، 0.8) فإنه يجب علينا أن نقوم برمي قطعة العملة المعدنية مائة مرة حتى توجد قيمة مشاهدة للمائة متغير عشوائي، ولذا، فإنه يمكننا أن نعدّ بدهيا، إلى حد ما، أن التقدير ماهو إلا القيمة المتحققة للمقدّر. وعموماً، فإن مقدراً \hat{P} هو دالة في المتغيرات العشوائية [انظر المعادلة (1B.18)]. لذا يجب أن يكون المقدّر كذلك متغيراً عشوائياً. فعلى سبيل المثال، يمكن أن يأخذ \hat{P} أي قيمة 0، 0.01، 0.02، ...، 0.99، 1.00، وذلك بناءً على نتائج رمي قطعة العملة المعدنية مائة مرة. وعليه، فإنه، إذا عددنا أن المائة رمية هذه للقطعة بوصفها تجربة عشوائية واحدة ذات عدد كبير من المفردات فإنه يستتج من ذلك أن P تكون متغيراً عشوائياً.

مقدّرات غير متحيّزة Unbiased estimators

يوصف المقدّر بأنه غير متحيّز إذا كانت قيمته المتوقعة أو وسطه الحسابي مساوياً للمعلمة التي نقوم بتقديرها. أي أنه إذا كان \hat{b} مقدراً لـ b فإن \hat{b} يكون غير متحيّز إذا كان:

$$E(\hat{b}) = b \quad (1B.19)$$

ومن جانب آخر، إذا كان $E(\hat{b}) \neq b$ يطلق على المقدّر \hat{b} بأنه مقدّر متحيّز للمعلمة b . وعلى سبيل المثال لتوضيح ذلك، افترض أن \hat{b} هو مقدّر للمعلمة P كما هو محدد في المعادلة (1B.18)، وباستخدام المعادلة رقم (1B.10) المرتبطة بالتوقعات، نجد أن:

$$\begin{aligned} E(\hat{P}) &= E\left(\frac{1}{100}X_1 + \frac{1}{100}X_2 + \cdots + \frac{1}{100}X_{100}\right) \\ &= \frac{1}{100}[E(X_1) + E(X_2) + \cdots + E(X_{100})] \end{aligned} \quad (1B.20)$$

ولما كانت دالة الاحتمال لكل X_i هي كما تظهر في المعادلة (1B.16) فإننا نجد أن:

$$E(X_i) = 0(1-P) + 1(P) = P \quad (1B.21)$$

والآن، بالتعويض من (1B.21) في (1B.20)، نحصل على:

$$E(\hat{P}) = \frac{1}{100}(100P) = P \quad (1B.22)$$

وهكذا، فقد أوضحنا أن \hat{P} هو مقدر غير متحيز للمعلمة P .

وعلى سبيل التوضيح للمقدر المتحيز بمثال، افترض أننا نريد تقدير قيمة P^* حيث إن $P^* = e^P$. في البداية، قد يبدو أن هذه الدالة ماهي الا تحويل مبسط للمشكلة التي نوقشت للتو، ولكن، لما كان P^* مرتبطا بالمعلمة P ارتباطا غير خطي فإن هذا التحويل يكون مهما.

إن المقدر الواضح لـ $P^* = e^P$ هو $\hat{P}^* = e^{\hat{P}}$ حيث \hat{P} معطاة في المعادلة (1B.18) ومن مناقشتنا السابقة يتضح أن:

$$E(\hat{P}^*) = E(e^{\hat{P}}) \neq e^{E(\hat{P})} = e^P = P^* \quad (1B.23)$$

وهكذا، فإن $E(\hat{P}^*) \neq P^*$ ، ولذا فإن \hat{P}^* مقدرا متحيزا للمعلمة P^* . وباختصار فإن كون P^* مقدرا غير متحيزا للمعلمة P لا يعني ضمنا أنه يمكننا استخدام \hat{P} مباشرة للحصول على مقدرات غير متحيزة للدوال غير الخطية في P . وكما سنرى في هذا الكتاب، فإن مشكلة الدوال غير الخطية مشكلة خطيرة من مشاكل الاقتصاد القياسي.

اتساق Consistency

مقدرنا الموضح بالمعادلة (1B.18) مبني على عينة عشوائية حجمها 100 مفردة، فإذا قمنا بدلا من ذلك برمي قطعة من العملات المعدنية عدد n من المرات فإنه يمكننا تعريف المقدر \hat{P} في الحال هذه على النحو التالي:

$$\hat{P}_n = \sum_{i=1}^n \frac{X_i}{n} \quad (1B.24)$$

وباستخدام (1B.10)، يمكن بسهولة إثبات أن $E(\hat{P}_n) = P$ (بمعنى أن \hat{P}_n هو مقدر غير متحيز) وباستخدام الصيغة الرياضية التي سوف نوجدتها في ملحق الفصل الثاني من هذا الكتاب، يمكننا أن نثبت أن تباين \hat{P}_n هو:

$$\sigma_{\hat{P}_n}^2 = \frac{1}{n^2} (\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2) \quad (1B.25)$$

حيث إن σ_i^2 هو تباين X_i الذي يمكن حسابه بالاستعانة بالدالة الاحتمالية لـ X_i والمعرفة في (1B.16) على النحو التالي:

$$\begin{aligned} \sigma_i^2 &= E[X_i - E(X_i)]^2 = E(X_i - P)^2 \\ &= [(0 - P)^2(1 - P) + (1 - P)^2 P] = P(1 - P) \end{aligned} \quad (1B.26)$$

وبالتعويض من المعادلة (1B.26) في المعادلة (1B.25) نحصل على:

$$\sigma_{\hat{P}_n}^2 = \frac{1}{n} [P(1 - P)] \quad (1B.27)$$

من المعادلة (1B.27)، يمكننا أن نرى أنه إذا اقترب حجم العينة من المالا نهائية ($n \rightarrow \infty$) فإن تباين \hat{P}_n يؤول إلى الصفر.* وتعني هذه النتيجة مع النتيجة التي تبين أن متوسط \hat{P}_n هو P أنه عندما تؤول n إلى مالا نهائية، فإن القيمة الأكثر

* يمكننا التعبير عن نفسه المعنى بعبارة بديلة وهي أنه إذا أصبح حجم العينة كبيرا بدرجة لانهائية، فإن تباين P_n سيكون مساويا للصفر.

احتمالا لـ \hat{P}_n تكون P . وبلغة فنية إحصائية يمكن اثبات أن احتمال اختلاف \hat{P}_n عن P بأي مقدار يقترب من الصفر كلما زاد حجم العينة، وأن هذا الاحتمال يؤول إلى الصفر في المالا نهاية. وباستخدام الرموز، تعني هذه العبارة مايلي:

$$\lim_{n \rightarrow \infty} \text{Prob}(|\hat{P}_n - P| > \varepsilon) = 0 \quad (1B.28)$$

حيث ε هي أي عدد صغير محدد سلفا. ^{الملاحظة البسيطة} وعندما يحقق أي مقدر شرطا مثل (1B.28) فإنه يوصف أنه مقدر متسق للمعلمة موضع الاهتمام. وهكذا فإن \hat{P}_n هو مقدر متسق لـ P . وبصفة عامة يمكن القول إن \hat{b} يكون مقدرًا متسقًا للمعلمة b إذا كان:

$$\lim_{n \rightarrow \infty} \text{Prob}(|\hat{b} - b| > \varepsilon) = 0 \quad (1B.29)$$

وغالبا ما يكتب شرط الاتساق (الموجود في 1B.29) مختصرا $P \lim \hat{b} = b$. ويعني هذا الشرط الأخير (مرة أخرى) أنه إذا آل حجم العينة إلى مالا نهاية فإن احتمال أن تأخذ \hat{b} قيمة أخرى غير قيمة b يساوي الصفر. وأخيرا، إذا كان \hat{c} مقدرًا غير متسق، حيثئذ فإن:

$$\lim_{n \rightarrow \infty} \text{Prob}(|\hat{c} - c| > \varepsilon) = 0 \quad (1B.30)$$

ومقدرا كهذا يوصف بعدم الاتساق.

دالة الكثافة المشتركة: إيضاحات

يوجد عديد من الحالات التي تتحدد فيها قيم أكثر من متغير عشوائي واحد نتيجة إجراء تجربة عشوائية. افترض، مثلا، القيام بتجربة يختار فيها الشخص عشوائيا، حيث يسجل فيها وزنه وطوله وعمره ونرمز لها بـ W ، H ، A على التوالي. في هذه الحال، نجد أن قيم ثلاثة متغيرات عشوائية تحدد جميعها بوساطة التجربة. ومثال آخر، افترض القيام بتجربة رمي قطعتي عملة معدنية عشوائيا، وافترض أن $X = 1$ إذا ظهرت الكتابة و $X = 0$ إذا ظهرت الشارة، في هذه الحال تحدد التجربة قيم

متغيرين عشوائيين. وهكذا، فإن مناقشتنا الأولية للاستقلال وعدمه تتضمن مثالا آخر. يجب أن يكون واضحا، عموما، أن التجربة الواحدة يمكن أن تحدد قيم عدد n من المتغيرات العشوائية، حيث تكون n عددا صحيحا موجبا.

افترض حال تحديد قيم متغيرين عشوائيين X و Y ولغرض التوضيح افترض أن قيم X الممكنة هي $x=1,2$ وقيم Y هي $y=1,2,3$ ، ويعني هذا أن هناك ستة أزواج ممكنة من القيم المناظرة لكل من X و Y حيث تظهر واحدة من القيم لكل من X ، Y عند كل تجربة، والأزواج الست المحتملة من القيم هي $(1,1)$ ، $(1,2)$ ، $(1,3)$ ، $(2,1)$ ، $(2,2)$ و $(2,3)$. حيث إن الحالة $(1,1)$ مثلا تناظر $X=1, Y=1$. والحالة $(1,2)$ تناظر $X=1, Y=2$ وهلم جرا. ولأكمال الصورة، افترض أن احتمالات الحصول على هذه الأزواج الست من القيم هي 0.1 ، 0.2 ، 0.15 ، 0.25 ، 0.1 و 0.2 على الترتيب، ويشير هذا، على سبيل المثال، إلى أن احتمال أن $(X=1, Y=2)$ هو 0.2 ، وهلم جرا.

وتعرف دالة الكثافة المشتركة (ويشار إليها، باختصار، بالكثافة المشتركة) لعدد من المتغيرات العشوائية بأنها الدالة التي تعطي الاحتمال الذي تأخذه كل مجموعة من المتغيرات العشوائية المناظرة لكل نتيجة محتملة من النتائج الممكنة، ويعني هذا للحالة السابقة أنه إذا كانت $f(X, Y)$ هي دالة الكثافة المشتركة لـ X و Y حيث إن $X=1,2$ ، $Y=1,2,3$ حينئذ فإن $f(1,3) =$ احتمال $(X=1, Y=3) = 0.15$ وهلم جرا.

ويمكن وصف دالة الكثافة المشتركة $f(x,y)$ لحالتنا التوضيحية هذه في شكل

الجدول التالي:

(x, y)	$(1,1)$	$(1,2)$	$(1,3)$	$(2,1)$	$(2,2)$	$(2,3)$
$f(x, y)$	0.1	0.2	0.15	0.25	0.1	0.2

(1B.31)

ويلاحظ من الجدول (1B.31) أن مجموع، الاحتمالات يساوي الواحد الصحيح والسبب في ذلك، أنه ينبغي أن تأخذ X و Y واحدة من الأزواج الست

المعطاة في الجدول. ولأغراض الرجوع إلى الأدب الإحصائي نلاحظ أن دالة الاحتمال المشترك الموجودة في (1B.31) يمكن التعبير عنها، غالباً بطريقة بديلة على النحو التالي:

x/y	1	2	3
1	0.1	0.2	0.15
2	0.25	0.1	0.2

(1B.32)

واتساقاً مع نقاشنا للمعادلة (1B.1)، نعرف دالة الاحتمال المشترك $f(x,y)$ المناظرة لأي زوج من القيم غير الممكنة لكل من X و Y بأنها تساوي الصفر. ومثال ذلك يتضح في الدوال التالية: $f(1, 1.5) = f(52.3, 2) = 0$ والسبب في هذا أن احتمال الحصول على نتيجة غير ممكنة يجب أن يكون صفراً. وعموماً يمكن تعريف قيمة الكثافة المشتركة المناظرة لمجموعة غير ممكنة من القيم للمتغيرات العشوائية الموجودة بالدالة بأنها مساوية للصفر.

وتحدد الكثافة المشتركة لـ X و Y جميع الاستدلالات الاحتمالية المرتبطة بـ X و Y . ولتوضيح ذلك، نستخدم (1B.31) أو (1B.32) لنجد أن احتمال أن يكون $(X=1, Y=1)$ أو $(X=2, Y=3)$ هو $(0.1+0.2=0.3)$. وبالمثل، فإن احتمال $(X=2, Y \leq 2)$ هو $0.25+0.1=0.35$.

فإذا افترضنا -زيادة في التوضيح- أننا مهتمون، فقط، بالمتغير X وخصوصاً احتمال أن يكون $X=1$ فإننا نرى من الجدول أن $X=1$ تناظر الحالات $(1,1)$ ، $(1,2)$ ، و $(1,3)$ ولا توجد هناك احتمالات أخرى. ولذلك، فإن احتمال $(X=1)$ = [احتمال أن $(X=1, Y=1)$ + احتمال $(X=1, Y=2)$ + احتمال $(X=1, Y=3)$] = $0.1+0.2+0.15=0.45$ ، وبالمثل، نجد أن احتمال $(X=2)$ يعادل $(0.25+0.1+0.2=0.55)$. وللإشارة المستقبلية نلاحظ أن احتمال $(X=1)$ يمكن الحصول عليه من دالة الكثافة المشتركة $f(x,y)$ عن طريق جمع $f(1,y)$ عبر كل القيم الممكنة لـ $y=1,2,3$.

عرفنا، من قبل، دالة الكثافة الاحتمالية لمتغير عشوائي، بأنها الدالة التي تعطي احتمالات أن يأخذ المتغير العشوائي كل قيمة من قيمه الممكنة. وفي المثال السابق، حددنا احتمال $(X=1) = 0.45$ واحتمال $(X=2) = 0.55$ ، ولاتوجد هناك قيم أخرى ممكنة لـ X . افترض الآن أن $g(x)$ هي دالة الكثافة الاحتمالية لـ X حيث إن $(x=1,2)$ ، مما يستتبع معه أن $g(1)=0.45$ و $g(2)=0.55$ أو في الشكل الجدولي التالي:

x	1	2
$g(x)$	0.45	0.55

(1B.33)

وبالمثل، إذا افترضنا أن $h(y)$ هي دالة الكثافة الاحتمالية للمتغير العشوائي Y حيث $(y=1,2,3)$ ، وبشكل متماثل، تماماً، لمسبق فإن قيم $h(y)$ تحدد من الجدول التالي: [كما حددناها من قبل لـ $g(x)$]:

y	1	2	3
$h(y)$	0.35	0.3	0.35

(1B.34)

وباختصار، فإنه يمكننا تحديد دالة الاحتمال لكل من X و Y من دالة الاحتمال المشترك لـ X و Y وسيتم ذلك بصورة رياضية Formally الآن.

دالة الكثافة المشتركة: تعميمات

افترض أن X و Y هي متغيرات عشوائية متقطعة، وأن قيمها الممكنة هي x_1, x_2, \dots, x_n و y_1, y_2, \dots, y_m وأن دالة الكثافة المشتركة لهما هي:

x, y	x_1, y_1	...	x_n, y_1	x_1, y_2	...	x_n, y_m
$f(x, y)$	$f(x_1, y_1)$...	$f(x_n, y_1)$	$f(x_1, y_2)$...	$f(x_n, y_m)$

(1B.35)

افترض - الآن - أن دالة الكثافة الاحتمالية لـ X هي $f_1(x)$ حيث $(x = x_1, \dots, x_n)$ ودالة الكثافة لـ Y هي $f_2(y)$ حيث $(y = y_1, \dots, y_m)$ في ضوء المناقشة السابقة، يتبين لنا أن:

$$\begin{aligned}
 f_1(x_1) &= f(x_1, y_1) + \cdots + f(x_1, y_m) \\
 f_1(x_2) &= f(x_2, y_1) + \cdots + f(x_2, y_m) \\
 &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\
 f_1(x_n) &= f(x_n, y_1) + \cdots + f(x_n, y_m)
 \end{aligned}
 \tag{1B.36}$$

المعادلة (1B.36)

وباستخدام رمز الجمع يمكننا أن نعبر عن (1B.36) على النحو التالي:

$$f_1(x) = \sum_{i=1}^m f(x, y_i), \quad x = x_1, \dots, x_n \tag{1B.37}$$

وتوجد علاقة مماثلة بين $f_2(y)$ و $f(x, y)$ وبالتحديد، تكون:

$$f_2(y) = \sum_{i=1}^n f(x_i, y), \quad y = y_1, \dots, y_m \tag{1B.38}$$

دالة الكثافة المشتركة: التوقعات

عرفنا من قبل القيمة المتوقعة لـ X على النحو التالي:

$$E(X) = x_1 f_1(x_1) + \cdots + x_n f_1(x_n) \tag{1B.39}$$

المعادلة

ويمكننا في ضوء المعادلات (1B.36) أن نحدد أيضا $E(X)$ بدلالة الاحتمال المشترك لكل من X و Y على النحو التالي:

$$\begin{aligned}
 E(X) &= x_1 [f(x_1, y_1) + \cdots + f(x_1, y_m)] \\
 &\quad + x_2 [f(x_2, y_1) + \cdots + f(x_2, y_m)] \\
 &\quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\
 &\quad + x_n [f(x_n, y_1) + \cdots + f(x_n, y_m)]
 \end{aligned}
 \tag{1B.40}$$

ومن الواضح أنه يمكن تحديد $E(Y)$ أيضا بدلالة $f_2(y)$ أو $f(x, y)$.

دوال الكثافة المشتركة: توقعات دوال المتغيرات العشوائية

افترض أن $h(x,y)$ هي دالة محدودة bounded لكل من X و Y ، ونعني بالقول بأن الدالة محدودة بأن $|h(x,y)|$ نهائية finite لكل قيم x ، $(x = x_1, \dots, x_n)$ ، وقيم y ، $(y = y_1, \dots, y_m)$. والقيم الممكنة للدالة $h(x,y)$ هي:

$$h(x_1, y_1), h(x_1, y_2), \dots, h(x_1, y_m), h(x_2, y_1), \dots, h(x_n, y_m)$$

نعرف القيمة المتوقعة للدالة $h(x,y)$ في هذه الحالة على النحو التالي:

$$\begin{aligned} E[h(X,Y)] &= h(x_1, y_1)f(x_1, y_1) + \dots + h(x_1, y_m)f(x_1, y_m) \\ &\quad + h(x_2, y_1)f(x_2, y_1) + \dots + h(x_2, y_m)f(x_2, y_m) \\ &\quad \vdots \\ &\quad + h(x_n, y_1)f(x_n, y_1) + \dots + h(x_n, y_m)f(x_n, y_m) \end{aligned} \quad (1B.41)$$

وتفسير $E[h(X,Y)]$ واضح ويتناسق مع حالة دوال الاحتمال ذات المتغير الواحد univariate case، وبالتحديد فإن القيمة المتوقعة للدالة $h(X,Y)$ تعرف بأنها المجموع المرجح لجميع قيمها الممكنة (حيث أن أوزان الترجيح هي الاحتمالات المناظرة للقيم).

وللتوضيح افترض الحالة الخاصة حيث إن $h(X,Y) = X$ ، حيث نجد من المعادلة (1B.41) أن:

$$\begin{aligned} E[h(X,Y)] &= x_1[f(x_1, y_1) + \dots + f(x_1, y_m)] \\ &\quad + x_2[f(x_2, y_1) + \dots + f(x_2, y_m)] \\ &\quad \vdots \\ &\quad + x_n[f(x_n, y_1) + \dots + f(x_n, y_m)] \end{aligned} \quad (1B.42)$$

وباستخدام المعادلة (1B.36)، يمكننا اختصار هذا إلى:

$$E[h(X,Y)] = x_1f_1(x_1) + x_2f_1(x_2) + \dots + x_nf_1(x_n) \quad (1B.43)$$

الذي يكون متماثلاً مع (1B.39)، وهذا يعني أن (1B.39) هي حالة خاصة من (1B.41).

توضيح: تغاير X و Y (Covariance of X and Y)

افترض أن X و Y هي متغيران عشوائيان مميز مستمران، لهما دالة كثافة مشتركة $f(x,y)$ ، حيث إن $(x = x_1, \dots, x_n)$ و $(y = y_1, \dots, y_m)$ ، افترض، أيضاً أن $E(X) = u_x$ وأن $E(Y) = u_y$ حيث إن $E(X)$ ، $E(Y)$ هي القيم المتوقعة لكل من X و Y على الترتيب. لذا، فإن تغاير X و Y (أو $\sigma_{x,y}$) يعرف على النحو التالي:

$$\sigma_{x,y} = [(X - u_x)(Y - u_y)] \quad (1B.44)$$

ويوجد في الفصل الثاني مناقشة وتفسير لمفهوم التغاير بين متغيرين عشوائيين.

ويمكن حساب تغاير X و Y ($\sigma_{x,y}$) باستخدام المعادلة (1B.41) وذلك عن طريق جعل $h(X,Y) = (X - u_x)(Y - u_y)$. وعلى سبيل المثال، افترض أن دالة الاحتمال المشترك لـ X و Y هي المعطاة في (1B.31)، حيث، باستخدام (1B.40)، فإننا نحصل على:

$$E(X) = 1[0.1 + 0.2 + 0.15] + 2[0.25 + 0.1 + 0.2] = 1.55 \quad (1B.45)$$

وبالمثل، فإن:

$$E(Y) = 1[0.1 + 0.25] + 2[0.2 + 0.1] + 3[0.15 + 0.2] = 2.00 \quad (1B.46)$$

وباستخدام (1B.41) نحصل على:

$$(1B.47)$$

$$\begin{aligned} \sigma_{x,y} &= E[(X - 1.55)(Y - 2.00)] \\ &= (1 - 1.55)(1 - 2)(0.1) + (1 - 1.55)(2 - 2)(0.2) + (1 - 1.55)(3 - 2)(0.15) \\ &\quad + (2 - 1.55)(1 - 2)(0.25) + (2 - 1.55)(2 - 2)(0.1) + (2 - 1.55)(3 - 2)(0.2) \\ &= -0.050 \end{aligned}$$

دوال الكثافة المشتركة: مناقشة أكثر عمومية

نحاول في هذا المبحث أن نعمم المفاهيم الأساسية التي حصلنا عليها أعلاه لحالة من ثلاثة متغيرات عشوائية، ومنها يمكن التعميم، أيضاً، مباشرة.

افترض وجود ثلاثة متغيرات عشوائية متقطعة X, Y, W قيمها الممكنة على التوالي $(x=x_1, \dots, x_n)$ ، $(y=y_1, \dots, y_m)$ و $(w=w_1, \dots, w_s)$ في هذه الحالة يكون لدينا n قيم ممكنة لـ X ، m قيم ممكنة لـ Y ، وأخيراً s قيم ممكنة للمتغير W .

افترض الآن أن $p(x, y, w)$ هي دالة الكثافة المشتركة لـ X, Y, W ، حيث على سبيل المثال $P(x_1, y_3, w_4) = \text{Prob}(X=x_1 \text{ و } Y=y_3, W=w_4)$ ، وحيث إن التوسع المباشر للمناقشة يتضمن التالي:

الملاحظة الأولى: مجموع كل القيم للدالة $P(x, y, w)$ (حيث يوجد عدد nms منها) يساوي الواحد.

الملاحظة الثانية: يمكن تحديد الاحتمالات المتعلقة باثنين، فقط، من المتغيرات العشوائية (X و Y ، مثلاً) من دالة الكثافة المشتركة للمتغيرات العشوائية الثلاثة (X, Y, W)، ولتوضيح هذه الملاحظة، دعنا نعود مرة ثانية لتوسيع النتيجة (1B.36) التي تعطينا:

$$\text{Prob}(X = x_i \text{ and } Y = y_j) = p(x_i, y_j, w_1) + p(x_i, y_j, w_2) + \dots + p(x_i, y_j, w_s) \quad (1B.48)$$

ويمكننا أن نعطي مثالين آخرين أكثر وضوحاً على النحو التالي:

$$\text{Prob}(X = 2 \text{ and } Y = 10) = p(2, 10, w_1) + p(2, 10, w_2) + \dots + p(2, 10, w_s) \quad (1B.49)$$

و

$$\text{Prob}(X = 3 \text{ and } Y = 7) = p(3, y_1, 7) + p(3, y_2, 7) + \dots + p(3, y_m, 7) \quad (1B.50)$$

مع ملاحظة أنه، في جميع الحالات، جمعت دالة الكثافة المشتركة لكل القيم الممكنة للمتغير الذي لم يظهر في دالة الاحتمال.

الملاحظة الثالثة: بافتراض أن دالة الكثافة المشتركة لـ X و Y هي $f(x, y)$ حيث إن $(x = x_1, \dots, x_n)$ و $(y = y_1, \dots, y_m)$ ، فإن نتيجة المعادلة (1B.48) تعطينا الدالة $f(x, y) -$ أي قيمة دالة الكثافة المشتركة المناظرة لـ y_i و x_i .

ولما كان من الممكن القيام بالحسابات الموجودة في (IB.48) لكل واحد من أزواج القيم: $(y_1, x_1), \dots, (y_m, x_1), (y_1, x_2), \dots, (y_m, x_2), \dots, (y_1, x_n), \dots, (y_m, x_n)$ فإنه يمكن تحديد دالة الكثافة المشتركة الكلية لكل من X و Y من دالة الكثافة المشتركة X و Y و W على النحو التالي:

$$\begin{aligned} f(x, y) &= p(x, y, w_1) + p(x, y, w_2) + \dots + p(x, y, w_s) \\ &= \sum_{i=1}^s p(x, y, w_i) \end{aligned} \quad (IB.51)$$

حيث تأخذ x أي قيمة من القيم x_1, \dots, x_n وأيضا تأخذ y أي قيمة من القيم y_1, \dots, y_m . وبنفس المنطق يمكن تحديد دالة الكثافة المشتركة لكل من W, X ، مثلاً $g(x, w)$ ودالة الكثافة المشتركة لـ Y و W ، مثلاً $h(y, w)$ باستخدام دالة الكثافة المشتركة لـ W, Y, X وذلك على النحو التالي:

$$\begin{aligned} g(x, w) &= p(x, y_1, w) + p(x, y_2, w) + \dots + p(x, y_m, w) \\ &= \sum_{i=1}^m p(x, y_i, w) \end{aligned} \quad (IB.52)$$

وأيضاً:

$$\begin{aligned} h(y, w) &= p(x_1, y, w) + p(x_2, y, w) + \dots + p(x_n, y, w) \\ &= \sum_{i=1}^n p(x_i, y, w) \end{aligned} \quad (IB.53)$$

ويلاحظ أنه، في جميع الحالات جمعت دالة الكثافة المشتركة على مدى جميع القيم الممكنة للمتغير الذي لا يناظر الكثافة في الجانب الأيسر من العلاقة. وتبين لنا مناقشتنا السابقة أن المعلومات المرتبطة بدالة الكثافة المشتركة لـ X, Y, W تتضمن، بدورها، المعلومات حول دوال الكثافة المشتركة لأي زوج من هذه المتغيرات. ولما كانت دالة الكثافة لـ X يمكن تحديدها من دالة الكثافة المشتركة لـ X, Y, W ،

و Y هلم جرا، فعليه يمكن القول أن دالة الكثافة المشتركة لـ Y, X و W تحدد أيضا كثافة X وكثافة Y وكثافة W .

نتيجة مهمة للاستقلال

لنأخذ، مرة أخرى، ثلاثة متغيرات عشوائية هي Y, X و W بدالة كثافة مشتركة $p(x, y, w)$ حيث إن $(x = x_1, \dots, x_n)$ ، $(y = y_1, \dots, y_m)$ و $(w = w_1, \dots, w_s)$ افترض مرة أخرى أيضا أن دالة الكثافة لـ X هي $f_1(x)$ ، ودالة الكثافة لـ Y هي $f_2(y)$ ، ودالة الكثافة لـ W هي $f_3(w)$. فإذا كانت Y, X و W مستقلة عن بعضها بعضا فإنه يمكن حساب دالة الكثافة المشتركة $p(x, y, w)$ على النحو التالي:

$$p(x, y, w) = f_1(x)f_2(y)f_3(w) \quad (1B.54)$$

ويعني ذلك أنه إذا كانت المتغيرات العشوائية Y, X و W مستقلة عن بعضها بعضا فإن دالة الكثافة المشتركة لها تساوي حاصل ضرب دوال الكثافة الفردية لكل منها. ويمكن أن نقوم بتعميم تلك النتيجة فنذكر أن دالة الكثافة المشتركة لأي عدد من المتغيرات العشوائية سوف يكون مساويا ضرب دوال الكثافة الفردية لكل منها إذا كانت تلك المتغيرات مستقلة عن بعضها بعضا.

ولما كانت $f_1(x_i) = \text{prob}(X = x_i)$ وهلم جرا. فإن أحد النتائج المهمة للمعادلة

(1B.54) هو أنه إذا كانت كل من Y, X و W مستقلة عن بعضها بعضا فإن:

$$\begin{aligned} \text{Prob}(X = x_i \text{ and } Y = y_j \text{ and } W = w_r) \\ = \text{Prob}(X = x_i) * \text{Prob}(Y = y_j) * \text{Prob}(W = w_r) \end{aligned} \quad (1B.55)$$

أي أن الاحتمال المشترك عند $X = x_i$ و $Y = y_j$ و $W = w_r$ يمكن حسابه عن طريق حاصل ضرب الاحتمالات الفردية في بعضها بعضا حيث لا يوجد تداخل بين المتغيرات العشوائية الثلاثة. وللتوضيح، فإن كل جزء يحدد دالة الاحتمال المشترك (1B.55) يرتبط فقط بالمتغير العشوائي المناظر له. وعموماً، لن يكون الحال كهذه إذا كانت المتغيرات مرتبطة ببعضها بعضا (أو غير مستقلة عن بعضها بعضا) وذلك حسب ما

رأينا من قبل عند مناقشتنا للاستقلال الإحصائي .

والآن، افترض أن $f(x, y)$ هي دالة كثافة مشتركة لـ Y, X ، فطالما أن:

$$f(x, y) = p(x, y, w_1) + p(x, y, w_2) + \dots + p(x, y, w_s)$$

وبافتراض الاستقلال، فإن المعادلة (1B.54) تتضمن مايلي:

(1B.56)

$$f(x, y) = f_1(x)f_2(y)[f_3(w_1) + f_3(w_2) + \dots + f_3(w_s)] = f_1(x)f_2(y)$$

ولما كان المقدار الذي يوجد بين القوسين [] هو مجموع الكثافة لـ w عبر كل القيم الممكنة لها، فإنه يكون، حينئذ، مساويا الواحد الصحيح. ومرة أخرى إذا افترضنا أن $h(y, w)$ و $g(x, w)$ هي دوال الكثافة المشتركة لكل من (w, Y) و (w, X) على الترتيب فإنه، بطريقة مشابهة للتي توصلنا بها إلى (1B.56)، يمكن أن نصل إلى النتائج التالية:

$$h(y, w) = f_2(y)f_3(w) \quad (1B.57)$$

وأیضا:

$$g(x, w) = f_1(x)f_3(w) \quad (1B.58)$$

وفي الحقيقة، تناظر النتائج الموجودة في المعادلتين السابقتين (1B.57) و (1B.58) النتيجة العامة التالية: افترض وجود متغيرات عشوائية عددها q هي X_1, X_2, \dots, X_q مستقلة عن بعضها البعض حيث يمكن اختصار كثافتها المشتركة بطريقة مشابهة لتلك الموجودة في (1B.54). عندئذ، فإن أي مجموعة جزئية من هذه المتغيرات العشوائية تكون مستقلة عن بعضها بعضا، كما تكون دالة الكثافة المشتركة لها مساوية حاصل ضرب دوال احتمال جميع متغيرات هذه المجموعة، فمثلا، إذا كانت X_1, X_2, \dots, X_{10} متغيرات مستقلة عن بعضها بعضا، فإن المتغيرات كانت X_1, X_3, X_7 تكون مستقلة عن بعضها بعضا كذلك، ويظهر ذلك في المعادلة التالية:

$$\text{Prob}(X_1 = 3 \text{ and } X_8 = 7) = \text{Prob}(X_1 = 3) * \text{Prob}(X_8 = 7) \quad (1B.59)$$

ومن الواضح أن خاصية الاستقلال الإحصائي المشترك للمتغيرات العشوائية تبسط كثيرا حساب الاحتمالات المشتركة.

تطبيق شروط الاستقلال

افترض أنه لدينا عملة معدنية، وأن احتمال ظهور الكتابة عند رمي هذه العملة هو P ، ولذا فإن احتمال ظهور الشارة T هو $1-P$. افترض أنه تم رمي هذه العملة عشوائيا n من المرات، دع $X_i = 1$ إذا اسفرت الرمية i عن ظهور الكتابة، $X_i = 0$ إذا أظهرت الرمية الشارة، وأن $(i = 1, \dots, n)$. ولما كان رمي العملة يتم عشوائيا، فإنه سيكون لدينا عدد n من المتغيرات العشوائية المستقلة عن بعضها بعضا وهي X_n, \dots, X_1 مع ملاحظة أن احتمال $p = (X_i = 1)$ وأن احتمال $1-p = (X_i = 0)$.

والآن، فإن احتمال أن تسفر الرميات الأولى وعددها s عن ظهور الكتابة وأن تسفر الرميات الأخيرة وعددها $n-s$ عن ظهور الشارة يمكن حسابه على النحو التالي: لما كانت X_n, \dots, X_1 هي متغيرات عشوائية مستقلة عن بعضها بعضا فإن هذا الاحتمال يكون [انظر (1B.55)]:

$$\begin{aligned} \text{Prob}(X_1 = 1 \text{ and } X_2 = 1 \text{ and } \dots \text{ and } X_s = 1 \text{ and } X_{s+1} = 0 \text{ and } \dots X_n = 0) \\ = \text{Prob}(X_1 = 1) \text{Prob}(X_2 = 1) \dots \text{Prob}(X_s = 1) \text{Prob}(X_{s+1} = 0) \dots \text{Prob}(X_n = 0) \quad (1B.60) \\ = P^s (1-P)^{n-s} \end{aligned}$$

افترض - الآن - أننا نريد حساب احتمال أن تسفر الرميات الأولى وعددها $n-s$ عن ظهور الشارة، وأن تسفر الرميات الأخيرة وعددها s عن ظهور الكتابة. يكون الاحتمال في هذه الحال هو الاحتمال نفسه في (1B.60) وذلك لأن:

المعادلة
السابقة

$$\begin{aligned}
& \text{Prob}(X_1 = 0 \text{ and } \dots \text{ and } X_{n-s} = 0 \text{ and } X_{n-s+1} = 1 \text{ and } \dots \text{ and } X_n = 1) \\
& = \text{Prob}(X_1 = 0) \dots \text{Prob}(X_{n-s} = 0) \dots \text{Prob}(X_{n-s+1} = 1) \dots \text{Prob}(X_n = 1) \quad (1B.61) \\
& = (1-p)^{n-s} P^s = P^s (1-P)^{n-s}
\end{aligned}$$

وللتعميم، افترض الآن، مجموعة معينة متسلسلة من الكتابة والشعار قد ظهرت حيث ظهر عدد s من الكتابة وعدد $n-s$ من الشارة. دع $P_{s,n}$ هي احتمال الحصول على هذه السلسلة، حيث يجب أن يكون واضحاً أن:

$$P_{s,n} = P^s (1-P)^{n-s} \quad (1B.62)$$

وسوف نحتاج هذه النتيجة فيما بعد.

أسئلة

(١) بين صحة المعادلة $\sum_{t=1}^n (X_t - \bar{X}) = 0$ عندما تكون $X_1=0$ ، $X_2=5$ ، $X_3=6$ ، و $X_4=1$.

(٢) أثبت أن:

$$\sum_{t=1}^n (aX_t + bY_t + cZ_t) = a \sum_{t=1}^n X_t + b \sum_{t=1}^n Y_t + c \sum_{t=1}^n Z_t$$

(٣) أثبت أن $\sum_{t=1}^n (X_t - \bar{X})(Y_t - \bar{Y})$ يمكن التعبير عنه على النحو

$$\cdot \sum_{t=1}^n X_t (Y_t - \bar{Y})$$

الفصل الثاني

نموذج انحدار المتغيرين

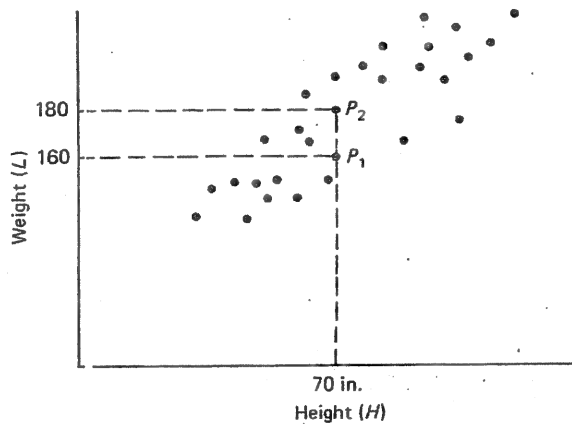
إحدى المشاكل الأساسية في الاقتصاد القياسي هي تطوير طرق فعالة لتقدير العلاقات الكمية بين المتغيرات الاقتصادية. وبدلالة المثال المعطى في الفصل الأول، فإن مانحتاجه هو طريقة ما تمكننا من الحصول على مقدرات للمعاملات a و b في دالة الاستهلاك حيث يمكن، من خلالها، التنبؤ بكيفية تغير الاستهلاك مع تغير مستوى الدخل المتاح بالإضافة إلى أشياء أخرى. في هذا الفصل، سوف نوضح المبادئ الأساسية لتقدير علاقة بين متغيرين. ونود أن نؤكد أن هذا الفصل هو أكثر الفصول أهمية في الكتاب. ففي هذا الفصل وفي القسم الأول من الفصل الثالث نقدم الهيكل الأساسي للمفاهيم الخاصة بالتقدير واختبار الفرضيات، أما المادة المعروضة في الفصول التالية (بما في ذلك، على سبيل المثال، تقدير العلاقات بين متغيرات عدة) فهي تعد أساساً، امتداداً مباشراً وبديهيًا لتحليلنا في حالة متغيرين.

(٢-١) قياس العلاقة الإحصائية بين متغيرين : التغير والارتباط

دعنا، في البداية نفترض أننا مهتمون، فقط، بوصف العلاقة الإحصائية بين متغيرين، ولا تتوفر لدينا أية فرضيات تتضمن أي نوع من العلاقات السببية بينهما. ومن ثم، فإن كل مانسعى إليه هنا هو تحديد ما إذا كان هناك أي نوع من الارتباط المنتظم بين المتغيرين.

وعلى سبيل المثال، افترض أننا سجلنا الوزن (L) بالرطل، والطول (H) بالبوصة لعينة من ٣٠ فرداً اختيروا عشوائياً. وكما فعلنا في الفصل الأول،

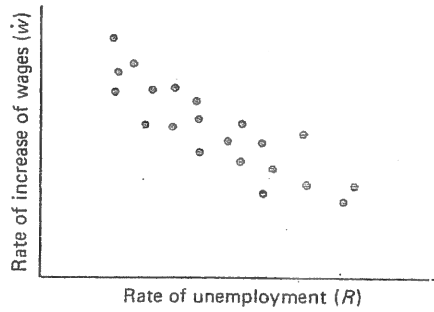
نستخدم شكل انتشار (١-٢) لتصوير المشاهدات. وكما في الفصل السابق، فإننا نلاحظ عدم وجود علاقة مؤكدة بين المتغيرين. فشخصان بالطول نفسه لن يكون لهما، عموماً، الوزن نفسه. ويمكن أن نرى في الشكل (١-٢) على سبيل المثال، أنه، بينما الشخصان الممثلان بالنقطتين P_1 و P_2 طول كل منهما ٧٠ بوصة إلا أن الأول يزن ١٦٠ رطلاً والثاني يزن ١٨٠ رطلاً. إلا أنه يظهر أن هناك علاقة من نوع ما بين الطول والوزن. فيبدو أن الأشخاص الأطول عادة ما يكونون ذوي وزن أكبر من القصار. ولذا، يبدو في المتوسط، أن هناك ارتباطاً موجباً بين الطول والوزن؛ فالقيم الأكبر من الوزن مقترنة على نحو منتظم بالقيم الأكبر من الطول.



شكل (١-٢)

وفي المقابل فإن شكل الانتشار (٢-٢) يوضح أن المتغيرين محل الاعتبار، وهما معدل التغير النسبي في الأجر (\dot{w}) ومعدل البطالة (R)، مرتبطان عكسياً. فارتفاع النقاط يتناقض كلما اتجهنا من اليسار إلى اليمين، الأمر الذي يوضح أن الزيادات الأسرع في الأجور مقترنة - بانتظام - بمعدلات بطالة أقل. وهذه قد لا تكون نتيجة مفاجئة بدرجة كبيرة. فعندما يكون الاقتصاد في حالة رواج، ويكون هناك أعداد قليلة من العمال العاطلين، فإننا نتوقع أن يقوم أصحاب العمل برفع الأجور بمعدلات أسرع نسبياً في محاولة منهم لزيادة الإنتاج كي يقابل

مستوى الطلب المرتفع على منتجاتهم. وعلى العكس من ذلك، عندما يكون الطلب الكلي منخفضا ونتيجة لذلك، البطالة مرتفعة، فسوف تكون هناك ضغوط أقل على الأجور لترتفع. وقد تصادف أن يكون المنحنى الممثل لنقاط هذا الانتشار معروفا بمنحنى فيليبس نسبة إلى A.W. Phillips وهو أول من لاحظ هذه العلاقة بين \dot{w} و R في بريطانيا.*



شكل (٢-٢)

التغاير

من بين الأسئلة التي نريد أن نسألها عن متغيرين : هل يرتبط هذان المتغيران بعلاقة طردية أم عكسية، أي هل القيم الأكبر لواحد منهما عادة ما تقترن بالقيم الأكبر للآخر (علاقة طردية)؟ أم أن القيم الأكبر للمتغير الأول عادة ما تكون مصاحبة للقيم الأصغر للثاني (علاقة عكسية)؟ وتسمى المعلمة التي ترصد هذه العلاقة الطردية أو العكسية «التغاير»، وبالنسبة لمتغيرين X و Y وسطاهما الحسابيان هما : $E(Y) = \mu_Y$ و $E(X) = \mu_X$ يعرف التغاير رياضيا :

* - انظر : A. W. Philips, "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957." *Economica* 25(Nov. 1958), pp. 283-299

$$\sigma_{X,Y} = E[(X - \mu_X)(Y - \mu_Y)] \quad (2.1)$$

أي أن التغير هو القيمة المتوقعة لحاصل ضرب $(Y - \mu_Y)$ في $(X - \mu_X)$ ، فإذا كان هذا التغير موجبا $\sigma_{X,Y} > 0$ فإنه يوضح أن القيم الأكبر من المتوسط الحسابي للمتغير X $[(X - \mu_X) > 0]$ تقترن، عادة، بقيم المتغير Y الأكبر من المتوسط $(Y - \mu_Y) > 0$ ، والعكس صحيح. وبديها فإن الفكرة هنا، وببساطة، هي أنه ليكون $\mu_{X,Y}$ موجبا فإن الحدين $(X - \mu_X)$ و $(Y - \mu_Y)$ لابد أن يكونا أما موجبين أو سالبين. ولهذا، نقول إذا كان X أكبر من قيمته المتوسطة μ_X فإن Y لابد أن يكون أكبر من قيمته المتوسطة μ_Y ، ومن ثم فإن X و Y ترتبطان طردياً. وعلى العكس من ذلك، إذا كان $\sigma_{X,Y} < 0$ فإن قيم X الأكبر من المتوسط سوف تكون عادة مصاحبة لقيم Y الأقل من المتوسط، الأمر الذي يشير إلى وجود علاقة عكسية بين المتغيرين.

أما عن الحالة الوسيطة، فهي، بالطبع، تلك التي يكون فيها $\sigma_{X,Y} = 0$. وفي هذه الحال فإن قيم X الأكبر من المتوسط سوف تكون مصاحبة بالقدر نفسه لكل من قيم Y الأصغر من المتوسط وقيمها الأكبر من المتوسط. ويوجد هناك حالتان لذلك، الأولى هي الحالة التي يكون فيها المتغيران مستقلين. فعلى سبيل المثال، إذا كان كل من X و Y مستقلاً:

$$\begin{aligned} \sigma_{X,Y} &= E[(X - \mu_X)(Y - \mu_Y)] \\ &= E[(X - \mu_X)E(Y - \mu_Y)] = 0 \end{aligned} \quad (2.2)$$

حيث :

$$E(X - \mu_X) = E(X) - E(\mu_X) = \mu_X - \mu_X = 0^{(*)}$$

والثانية هي الحالة التي يكون فيها المتغيران على علاقة ببعضهما بطريقة غير خطية،

* على القارئ أن يتذكر (كما هو مبين في الملحق B بالفصل الأول) أنه إذا كان المتغيران مستقلان فإن القيمة المتوقعة لحاصل ضربهما تساوي حاصل ضرب قيمتهما التوقعيتين. ويلاحظ أيضاً أن متوسط X أي μ_X

ثابت ولهذا فإن $E(\mu_X) = \mu_X$.

وسوف نرى مثالا على ذلك فيما بعد، ولكن، يتعين، عند هذه النقطة ملاحظة أنه، بسبب هذه الحالة غير الخطية، فإن التغير بين متغيرين إذا كان مساويا للصفر فإن هذا لا يتضمن أنهما مستقلان. وبدلا من ذلك فإن التغير المساوي للصفر يوحي، فقط، بأن المتغيرين لا تجمع بينهما علاقة خطية.

مقدر التغير

عملياً، لا يمكن أن نعرف، عموماً، قيمة $\sigma_{X,Y}$. وفي العادة، يتوافر لدينا، فقط، عينة عشوائية من القيم المشاهدة لكل من X و Y . وكما سبق، قد يتوافر لدينا، على سبيل المثال، الطول والوزن لعدد معين n ، من الأفراد اختيروا عشوائياً. وفي مثل هذه الحال، نقول إن لدينا عينة حجمها n عن L و H . وافترض الآن أن لدينا عينة حجمها n عن X و Y . وما نحتاجه هو طريقة ما لتقدير قيمة $\sigma_{X,Y}$ من هذه العينة من المشاهدات. وحيث إن $\sigma_{X,Y}$ تعرف بأنها القيمة المتوقعة لحاصل ضرب انحرافات المتغيرين عن وسطيهما الحسابيين $E[(X - \mu_X)(Y - \mu_Y)]$ فإن الطريقة الواضحة لتقدير $\sigma_{X,Y}$ هي حساب متوسط حاصل ضرب انحرافات X و Y عن وسطيهما الحسابيين بالعينة. وبأسلوب رياضي، افترض أن القيم المشاهدة n لكل من X و Y بالعينة هي X_1, \dots, X_n و Y_1, \dots, Y_n ومن ثم فإن مقدر التغير بين X و Y هو :

$$\hat{\sigma}_{X,Y} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n-1} \quad (2.3)$$

حيث :

* فعلى سبيل المثال X_5 هي القيمة المشاهدة الخامسة للمتغير X . وفي مثالنا الأسبق حيث H هي طول الشخص، L هي وزنه، فإن H_5 هي طول الشخص الخامس المشاهد، L_5 هي وزن الشخص نفسه.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \text{ و } \bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}$$

وسنوضح، لاحقاً، لماذا تقسم المعادلة (2.3) على (n-1) بدلاً من n. وعند هذه النقطة، على أي حال، نلاحظ أنه، طالما $(Y_i - \bar{Y})(X_i - \bar{X})$ هو حاصل ضرب الانحرافات رقم t بالعينة، فإن $\hat{\sigma}_{X,Y}$ ببساطة هو متوسط حواصل الضرب.* ولعله من المفيد أن نتوقف قليلاً عند هذه النقطة ونوضح المقصود بالرموز المستخدمة. ففي هذا الكتاب سوف نستخدم العلامة $\sigma_{X,Y}$ فوق متغير أو معلمة لنشير إلى «مقدر»، ومن ثم فإن $\hat{\sigma}_{X,Y}$ هي مقدر $\sigma_{X,Y}$ ، ويدل الجانب الأيمن من المعادلة (2.3) أن $\hat{\sigma}_{X,Y}$ مبني على أساس مجموع حاصل ضرب $(Y_i - \bar{Y})(X_i - \bar{X})$ لكل العينة $t=1, \dots, n$. ولتبسيط الرموز، فمن الآن وصاعداً لن نتكبد جهد كتابة $\sum_{i=1}^n$ ولكن نستخدم ببساطة Σ لتدل على أن عملية الجمع شاملة لجميع مشاهدات العينة، مالم نشر إلى غير ذلك. عدم تحيز $\hat{\sigma}_{X,Y}$

لقد عرفنا أن $\hat{\sigma}_{X,Y}$ هو مقدر التباين بين X و Y، علاوة على ذلك يمكن توضيح أن $E[\hat{\sigma}_{X,Y}] = \sigma_{X,Y}$ ، أي أن $\hat{\sigma}_{X,Y}$ مقدر غير متحيز لـ $\sigma_{X,Y}$ ، والفكرة هنا هي طالما أن $\hat{\sigma}_{X,Y}$ مستخرج من عينة عشوائية للمتغيرين X و Y، فإنه يعتبر هو نفسه متغيراً عشوائياً تتغير قيمته من عينة لأخرى. فعلى سبيل المثال، لو أن الوزن والطول لعدد 30 شخصاً (اختيروا عشوائياً) يعتمدان على الأشخاص المعنيين الذين اختيروا، فإن قيمة مقدر التباين بين L و H سوف تعتمد أيضاً على من

* لو أن الشخص الخامس المشاهد في عينتنا كان أطول بمقدار 3 بوصات وأثقل في الوزن بمقدار 15 رطلاً عن متوسطي العينة للأطوال والأوزان على التوالي، عندئذ $(L_5 - \bar{L})(H_5 - \bar{H}) = 45$. والمقدر $\hat{\sigma}_{H,L}$ لتباين L و H هو، ببساطة، متوسط حاصل الضرب للانحرافين السابقين لكل قيم العينة.

اختيروا. ويتبع ذلك أن قيمة هذا المقدّر سوف تتغير من عينة لأخرى. وبتعبير أدق فإن للمقدّر دالة احتمال تسمى توزيع المعاينة Sampling distribution. وتوحي النتيجة المشار إليها سابقا وهي $E[\hat{\sigma}_{X,Y}] = \sigma_{X,Y}$ أن متوسط توزيع المعاينة الخاص بالمقدّر هو قيمة المعلمة $\sigma_{X,Y}$.

وعلى الرغم من أن اثباتا رياضيا يعد خارج مجال هذا الكتاب، فإن مثالا قد يوضح بديهية أكثر ماذا نعني بقولنا أن $\hat{\sigma}_{X,Y}$ مقدّر غير متحيز لـ $\sigma_{X,Y}$ *. ولتفصيل القول السابق، افترض أننا أخذنا عدد M من العينات تتكون كل واحدة منها من ٣٠ فردا، ثم قسنا الوزن L والطول H لأفراد كل عينة. وعندئذ، يمكننا حساب $\hat{\sigma}_{L,H}$ لكل عينة بحيث يتوافر لدينا عدد M من القيم المقدرة للتغاير بين L, H . ويلاحظ أن قيم $\hat{\sigma}_{L,H}$ المتعددة سوف تكون عموما مختلفة. وعلى وجه التحديد فإننا نتوقع أن تكون بعض قيمنا المقدرة أكبر من $\sigma_{L,H}$ ، وبعضها أقل. والآن تذكر أن القيمة المتوقعة لـ $\hat{\sigma}_{L,H}$ هي $\sigma_{L,H}$. وهذا يتضمن أننا لو أخذنا متوسط للعدد M من القيم المقدرة، فإننا نتوقع أن تكون قيمة هذا المتوسط هي قيمة المعلمة $\sigma_{L,H}$ ، وبصورة رياضية أكثر، افترض أن $\hat{\sigma}_{L,H}$ هي هذا المتوسط :

$$\bar{\hat{\sigma}}_{L,H} = \sum_{i=1}^M \frac{\hat{\sigma}_{L,H_i}}{M}, \quad (2.4)$$

حيث إن $\hat{\sigma}_{L,H_i}$ هي مقدّر $\sigma_{L,H}$ للعينة i . ومن ثم فإن $E(\bar{\hat{\sigma}}_{L,H}) = \sigma_{L,H}$ ، بالإضافة إلى ذلك فإنه يمكن توضيح أنه تحت شروط عامة - لو كان العدد M من العينات لانهائيا فإن احتمال أن يختلف $\bar{\hat{\sigma}}_{L,H}$ عن $\sigma_{L,H}$ بأي مقدار مهما كان صغيرا سوف يساوي الصفر.

* الأمثلة التالية ليست «تعريفات بديهية» لمفهوم عدم التحيز، وإنما هي عروض بديهية لبعض النتائج التي يتضمنها عدم التحيز تحت شروط عامة.

وتفسر هذه النتيجة مباشرة، إذ في الواقع، عادة، ما يكون لدينا عينة واحدة، وعلى أساس هذه العينة، نحسب قيمة مقدرة للتغاير باستخدام الصيغة العامة المعطاة في (2.3). ولو أن هذه العينة قد اختيرت عشوائياً فإنها يمكن أن تكون أي واحدة من عدد لانتهائي للعينات (مثلاً، أي واحدة من العينات M). وعلى الرغم من ذلك، فإنه، بسبب النتيجة الوسيطة التي حصلنا عليها سابقاً، فلا يوجد سبب يدعو للاعتقاد بأن القيمة المقدرة من العينة المختارة سوف تفوق أو تقل عن القيمة المقابلة لمعلمة التغاير. وفي المقابل، لو أن $E(\hat{\sigma}_{X,Y}) = \sigma_{X,Y} + 5$ فإننا نتوقع أن تكون القيمة المحسوبة أعلى من $\sigma_{X,Y}$. ويمكن أخذ هذا التحيز بالاعتبار بأخذ $(\hat{\sigma}_{X,Y} - 5)$ كمقدر لـ $\sigma_{X,Y}$.

قد يكون هناك نوع من اللبس لدى القارئ لكون مقام المعادلة (2.3) هو $(n-1)$ بدلاً من حجم العينة بالكامل n . فعادة، عند حساب متوسط ما، فإننا نقسم مجموع القيم على عددها. وفي هذه الحال، بالرغم من وجود عدد n من الحدود الممثلة في مجموع البسط، فإن هذه الحدود n يمكن تخفيضها إلى $(n-1)$ حد لها المجموع نفسه. وبمعنى آخر، هناك $(n-1)$ معلومة فقط. والسبب في ذلك هو أن كلا من \bar{X} و \bar{Y} متوسطي العينة يظهران في البسط مع القيم المشاهدة للمتغيرين Y و X . وبديهيًا حتى نكون الحدود الـ $(n-1)$ الأولى في (2.3) فلا بد من معرفة: X_1, X_2, \dots, X_{n-1} و Y_1, Y_2, \dots, Y_{n-1} بالإضافة إلى \bar{X} و \bar{Y} حيث إن:

$$\bar{X} = \frac{\sum (Y_1 + \dots + X_n)}{n} \text{ و } \bar{Y} = \frac{\sum (Y_1 + \dots + X_n)}{n}$$

وبهذه المعلومات، يمكننا تحديد قيمتي كل من X_n و Y_n . وفي ظل هذه الظروف، فإن الحد الأخير في (2.3) أي $(Y_n - \bar{Y})(X_n - \bar{X})$ لا يتضمن أية معلومات جديدة. ونتيجة لذلك، فنحن نعرف مسبقاً أو يمكننا أن نحسب قيمة الحد الأخير في المجموع من المعلومات التي يحتوى عليها العدد $(n-1)$ الأول من الحدود. ويوصف هذا الشرط، غالباً، بالقول إن البسط في (2.3) له $(n-1)$ درجات حرية بما

معناه أن هنالك (n-1) معلومة مستقلة فقط . وحيث إن البسط له (n-1) درجات حرية، فقط، يمكن إثبات أن:

$$E[\Sigma(X_t - \bar{X})(Y_t - \bar{Y})] = (n-1)\sigma_{X,Y} \quad (2.5)$$

ونتيجة لذلك، فإن القسمة على (n-1) تجعل $\hat{\sigma}_{X,Y}$ مقدراً غير متحيز لـ $\sigma_{X,Y}$.

اتساق $\hat{\sigma}_{X,Y}$

من المفيد هنا، وللتحليل فيما بعد، أن نتبين خصائص $\hat{\sigma}_{X,Y}$ في حالة العينة الكبيرة. ونقصد بذلك سلوك $\hat{\sigma}_{X,Y}$ عندما يكبر حجم العينة إلى ما لانهاية. ولتوضيح ذلك، دعنا نأخذ، مثلاً، متوسط العينة \bar{X} لعدد من قيم X التي اختيرت عشوائياً. ونعلم من مبادئ الإحصاء أن متوسط \bar{X} هو μ_X وأن تباين X هو σ_X^2 / n حيث μ_X و σ_X^2 يشيران إلى متوسط المتغير العشوائي X وتباينه على التوالي، كما تشير n إلى حجم العينة. وإذا زاد حجم العينة n باستمرار، نرى أن تباين \bar{X} أي σ_X^2 / n يصبح، دائماً أقل، ويؤول إلى الصفر عندما تزداد n إلى ما لانهاية. والفكرة هنا هي أنه كلما أصبح حجم العينة أكبر فإن احتمال وقوع متوسط العينة \bar{X} داخل مدى محدد من متوسط المجتمع μ_X يزداد باستمرار. وعندما يصبح حجم العينة لانهاياً فإن تباين \bar{X} يصبح مساوياً للصفر وبالتالي فإن احتمال أن يساوي \bar{X} أي شيء آخر غير μ_X يصبح صفراً، ولهذا السبب تعد \bar{X} مقدراً متسقاً لـ μ_X .

عموماً، (كما هو موضح في الملحق B للفصل الأول) تعني خاصية الاتساق، أنه، في حالة كون حجم العينة لانهاياً، فإن احتمال أن يأخذ المقدّر قيمة تختلف بأي مقدار عن المعلمة المقابلة يساوي صفراً، فلو أن مقدراً (مثل \bar{X}) يتصف بالاتساق فإن هذا المقدّر يقال عنه إنه يتقارب في الاحتمال لمعلمته المقابلة (μ_X). ومن السهل أن نرى، في الأقل، بديهياً، أنه في حالة المعادلة (2.3) فإن $\hat{\sigma}_{X,Y}$ مقدّر متسق لـ $\sigma_{X,Y}$. فعندما يصبح حجم العينة لانهاياً فإن \bar{X} و \bar{Y}

تتقاربان في الاحتمال للمعلمتين μ_X و μ_Y على التوالي . ولهذا، فإن $\hat{\sigma}_{X,Y}$ يصبح متوسط العينة للقيم $(Y - \mu_Y)(X - \mu_X)$ المشاهدة في عينة ذات حجم لانهائي .
وحيث إن $E(X - \mu_X)(Y - \mu_Y) = \sigma_{X,Y}$ فيتضح أنه، تحت شروط عامة، إذا كانت العينة ذات حجم لانهائي فإن :

$$\hat{\sigma}_{X,Y} = \sigma_{X,Y}$$

وذلك باحتمال يساوي الواحد الصحيح* .

تفسير لـ $\hat{\sigma}_{X,Y}$

دعنا الآن نفسر $\hat{\sigma}_{X,Y}$ بدلالة شكل الانتشار (٢-٣)، لقد اوضحنا متوسطي العينة للقيم المشاهدة Y, X بالخطوط المنقطة، كما استخدمت هذه الخطوط لتقسيم الشكل رقم (٢-٣) إلى أربعة أقسام :

في القسم الأول: $(Y - \bar{Y}) > 0$ و $(X - \bar{X}) > 0$,

وعليه، فإن $(X - \bar{X})(Y - \bar{Y}) > 0$.

في القسم الثاني: $(Y - \bar{Y}) < 0$ و $(X - \bar{X}) > 0$,

وعليه فإن $(X - \bar{X})(Y - \bar{Y}) < 0$.

في القسم الثالث: $(Y - \bar{Y}) < 0$ و $(X - \bar{X}) < 0$,

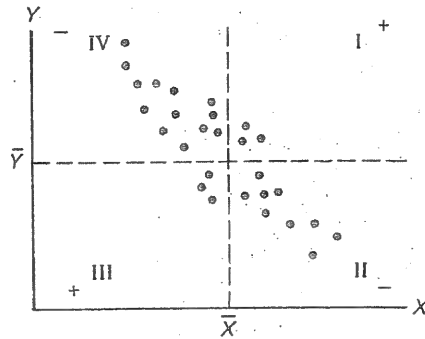
وعليه فإن $(X - \bar{X})(Y - \bar{Y}) > 0$.

في القسم الرابع: $(Y - \bar{Y}) > 0$ و $(X - \bar{X}) < 0$,

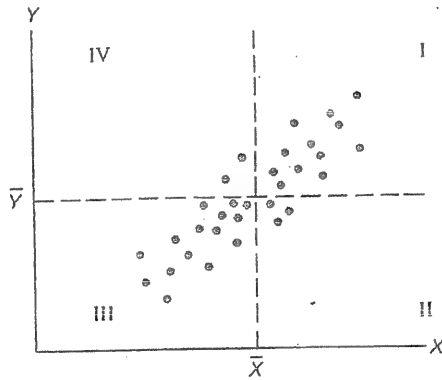
وعليه فإن $(X - \bar{X})(Y - \bar{Y}) < 0$.

* ملحوظة للقراء الأكثر دراية بالإحصاء يوجد هناك أشكال أخرى للتقارب بالإضافة إلى خاصية الاتساق التي ناقشناها في الملحق B في الفصل الأول، وأحد هذه الأشكال يسمى «بالتقارب مع احتمال 1». ونحن لانشير لهذا الشكل من التقارب بعبارتنا السابقة، ولكن بدلا من ذلك نحاول تبسيط الفكرة وجعل المادة بديهية، ولذا نصف $\lim_{n \rightarrow \infty} \text{prob}(|\hat{\sigma}_{X,Y} - \sigma_{X,Y}| > \varepsilon) = 0$. ولفظيا "إذا كان حجم العينة لانهائيا فإن

$\hat{\sigma}_{X,Y}$ سوف تساوي $\sigma_{X,Y}$ باحتمال 1".



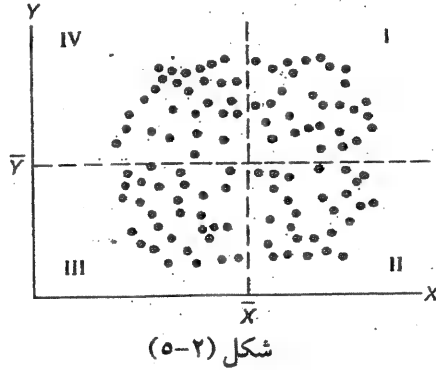
شكل رقم (٣-٢)



شكل رقم (٤-٢)

وتشير المشاهدات في الشكل (٣-٢) إلى وجود علاقة عكسية بين المتغيرين، ويتخذ هذا شكلاً تتركز فيه النقاط بالقسمين الثاني والرابع مع عدد قليل نسبياً من المشاهدات تقع في القسمين الأول والثالث. وحيث إن $(Y - \bar{Y})(X - \bar{X})$ سالبة في القسمين الثاني والرابع وموجبة في القسمين الأول والثالث فمن المتوقع أن تكون $\hat{\sigma}_{X,Y}$ سالبة في هذه الحالة. وبمعنى آخر فإن متوسط حاصل ضرب الانحرافات $(Y - \bar{Y})(X - \bar{X})$ سوف يكون سالباً. ومن ناحية أخرى، لو أن شكل الانتشار أظهر تركيزاً في القسمين الأول والثالث كما في الشكل (٤-٢) فمن المتوقع أن تكون $\hat{\sigma}_{X,Y}$ موجبة بالطريقة نفسها.

دعنا نأخذ الآن الحالة التي يكون فيها X و Y مستقلين، حيث لا يوجد هناك اقتران بينهما، فقيمة عالية للمتغير Y يكون احتمال اقترانها بقيمة عالية للمتغير X هو نفسه احتمال اقترانها بقيمة منخفضة للمتغير نفسه. وفي مثل هذه الحال، لا يتوقع أن يظهر شكل الانتشار بين X و Y اتجاها تصاعديا أو تنازليا. ويوضح الشكل (٢-٥) مثل هذه العلاقة، ففيه، نرى النقاط تتوزع بالتساوي تقريبا بين الأقسام الأربعة. ونتيجة لذلك، فإن القيم الموجبة لـ $(Y - \bar{Y})(X - \bar{X})$ الناجمة عن نقاط الانتشار بالقسمين الأول والثالث تلغي القيم السالبة المتولدة عن نقاط الانتشار بالقسمين الثاني والرابع ومن ثم، فإن القيمة المحسوبة لـ $\hat{\sigma}_{X,Y}$ تميل لأن تكون قريبة من الصفر.

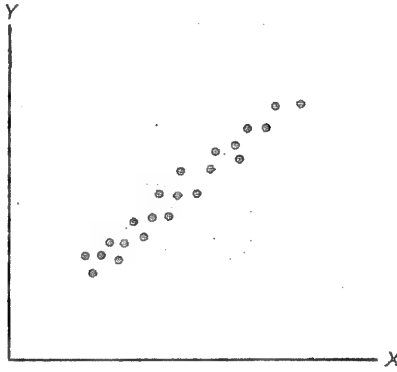


معامل الارتباط Correlation coefficient

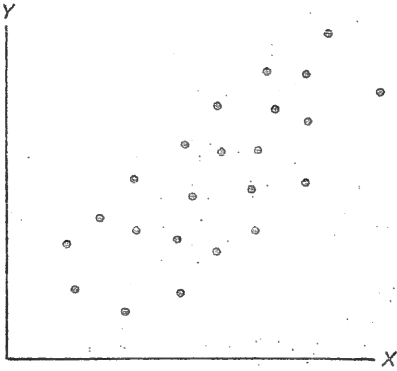
بالإضافة إلى معرفة ما إذا كان متغيران ما تجمعهما علاقة طردية أم عكسية نريد بوجه عام مؤشرا عن مدى قوة هذه العلاقة. والشكلان رقما (٢-٦)، (٢-٧) يقدمان على سبيل المثال حالتين للعلاقة الطردية بين X و Y وعلى الرغم من ذلك، فإن الاقتران الطردي في الشكل الأسبق أقوى، إلى حد ما، منه بالشكل الذي يليه. ويتحديد أكثر، يمكننا أن نرى أنه، بمعرفة قيمة X فحسب فإن التباين في Y بالشكل (٢-٦) أصغر نسبيا منه مقارنة بالشكل (٢-٧)*، ومن ثم فإن من

* يمكن للقارئ الأكثر معرفة اثبات أنه، في ظل تحقق شروط معينة، فإن تباين Y المشروط بمعرفة X ، يتغير عكسيا مع مربع معامل الارتباط.

المرغوب فيه أن يكون لدينا مقياس لهذه الخاصية للعلاقة بين X و Y . ولسوء الحظ، فإن مقياس التغير غير مناسب لتوضيح مدى قوة الاقتران لأن قيمته تعتمد على وحدات القياس المعينة للمتغيرات، فعلى سبيل المثال، سوف يكون التغير بين الطول والوزن أكبر لو أننا استخدمنا البوصة والأوقية في القياس بدلا من القدم والرطل.*



الشكل (٦-٢)



الشكل (٧-٢)

* افترض، على سبيل المثال، أن وحدة القياس هي $Z = aX$ بدلا من X حيث a ثابت، ومن ثم سوف يكون لدينا $E(Z - \mu_Z)(Y - \mu_Y) = E(aX - a\mu_X)(Y - \mu_Y) = aE(X - \mu_X)(Y - \mu_Y) = a\sigma_{X,Y}$ ومن ثم، $\sigma_{Z,Y} \neq \sigma_{X,Y}$.

ولكي نحصل على قياس ذي معنى لقوة الاقتران بين متغيرين، فنحن في حاجة إلى معلمة تكون مستقلة القيمة عن وحدات القياس. ويمكن اشتقاق هذه المعلمة بقسمة تغاير X و Y على الانحرافين المعياريين للمتغيرين. وبدقة، فإن قوة العلاقة بين X و Y يمكن توضيحها بمعامل الارتباط $\rho_{X,Y}$:

$$\rho_{X,Y} = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y} \quad (2.6)$$

حيث σ_Y, σ_X هما الانحرافان المعياريان لكل من Y, X .

$$\sigma_X = +\sqrt{E(X - \mu_X)^2} \text{ و } \sigma_Y = +\sqrt{E(Y - \mu_Y)^2}$$

نعود الآن إلى خصائص معامل الارتباط*. لاحظ أولاً أنه يأخذ دائماً إشارة التغاير نفسها، فطالما أن مقام المعادلة (2.6) موجب دائماً فإنه يتبع ذلك أن إشارة $\rho_{X,Y}$ سوف تكون هي نفسها إشارة البسط والذي هو التغاير بين المتغيرين. ومن ثم لو أن Y, X مرتبطان طردياً فإنه يتبع ذلك أن تكون $\rho_{X,Y} > 0$ ، ولو أنهما مرتبطان عكسياً فإن $\rho_{X,Y} < 0$. وفي هذه الحالة التي يكون فيها X و Y مستقلان فإن $\rho_{X,Y} = 0$ طالما أن $\sigma_{X,Y} = 0$ ، ومن ثم، فإن لمعامل الارتباط جميع خصائص التغاير من حيث توضيحه نوع العلاقة التي توجد بين المتغيرين.

ولكن، على العكس من التغاير فإن معامل الارتباط يوجد له مدى تتقلب قيمه بين حديه. وبتحديد أكثر فإن قيمة $\rho_{X,Y}$ لا بد أن تقع عند ناقص أو زائد واحد أو بينهما. إضافة إلى ذلك، كلما اقتربت $\rho_{X,Y}$ من الواحد في أي اتجاه كلما أصبح الارتباط الخطي أقوى بين المتغيرين (إما طردياً أو عكسياً)، وكلما اقتربت $\rho_{X,Y}$ من الصفر كلما كانت العلاقة أضعف. وتمثل $\rho_{X,Y} = 0$ حالة عدم وجود أي ارتباط بين المتغيرين. وبالنسبة للشكلين رقمهما (٢-٦) و (٢-٧) على سبيل المثال،

* ولو أننا غيرنا وحدة القياس كما هو في الحاشية السابقة باستخدام $Z = aX$ بدلا من X فإن معلمتنا $\rho_{X,Y}$ (على عكس $\rho_{X,Y}$) لن تتأثر. أي أن، $\rho_{Z,Y} = \rho_{X,Y}$ ونحن نترك إثبات ذلك تمرينا للقارئ.

فإنه من المتوقع أن يكون معامل الارتباط في حالة الشكل رقم (٢-٦) أكبر منه في حالة الشكل رقم (٢-٧)، وفي كلتا الحالتين فإنه موجب بالطبع.

والآن نثبت أنه إذا كانت X و Y مرتبطتين ارتباطاً تاماً وخطياً فإن $\rho_{X,Y}$ سوف تكون قيمته إما زائد واحد أو ناقص واحد، ولنرى ذلك*، دعنا نأخذ العلاقة التامة التالية:

$$Y = a + bX \quad (2.7)$$

وفي هذا المثال فإن نقاط شكل الانتشار سوف تقع جميعها على خط مستقيم ميله b وقاطعه a ، ومعامل الارتباط بين X و Y هو:

$$\rho_{X,Y} = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y}$$

وسوف نثبت الآن أن $\rho_{X,Y} = +1$ إذا كانت $b > 0$ وذلك بالتعبير عن كل من

σ_Y , $\sigma_{X,Y}$ بدلالة σ_X ونشتق أولاً μ_Y :

$$E(Y) = E(a + bX) = a + bE(X) = a + b\mu_X = \mu_Y \quad (2.8)$$

ولذا فإن التباين:

$$\begin{aligned} \sigma_{X,Y} &= E[(Y - \mu_Y)(X - \mu_X)] \\ &= E[(a + bX - a - b\mu_X)(X - \mu_X)] \\ &= E[b(X - \mu_X)^2] = b\sigma_X^2 \end{aligned} \quad (2.9)$$

وبالتعريف فإن تباين Y هو:

$$\sigma_Y^2 = E[(Y - \mu_Y)^2] \quad (2.10)$$

* المثال التالي يوضح أن $\rho_{X,Y}$ تساوي زائد واحد أو ناقص واحد لو أن X و Y مرتبطان ارتباطاً تاماً وخطياً. ولسوء الحظ، فإن إثبات $\rho_{X,Y}$ أن لا يمكن أن يفوق القيمة المطلقة للواحد تحت أي ظرف يقع خارج مجال هذا الكتاب (بالرغم من أنه قد يبدو معقولاً بديهياً).

ويمكن التعبير عن هذا التباين كمايلي :

$$\begin{aligned}\sigma_Y^2 &= E\left[(Y - \mu_Y)^2\right] = E\left[(a + bX - a - b\mu_X)^2\right] \\ &= E\left[b^2(X - \mu_X)^2\right] = b^2\sigma_X^2\end{aligned}\quad (2.11)$$

والانحراف المعياري للمتغير Y هو الجذر التربيعي الموجب $\sigma_Y = b\sigma_X$ ، ويمكن الآن تحديد معامل الارتباط :

$$\rho_{X,Y} = \frac{\sigma_{X,Y}}{\sigma_X\sigma_Y} = \frac{b\sigma_X^2}{(b\sigma_X)\sigma_X} \quad (2.12)$$

وسوف نترك للقارئ أن يثبت أنه إذا كانت $b < 0$ فإن $\rho_{X,Y} = -1$. (ملحوظة : إذا كانت $b < 0$ ، $\sigma_Y = -b\sigma_X > 0$).

وباختصار ، فإن معامل الارتباط يوضح كلا من اشارة العلاقة الخطية وقوتها بين متغيرين . والقيم الموجبة والسالبة لـ $\rho_{X,Y}$ توضح العلاقات الطردية والعكسية على التوالي ، وكلما اقترب $\rho_{X,Y}$ من زائد واحد أو ناقص واحد كلما زادت قوة العلاقة الخطية ، أي كلما زادت قوة الارتباط بين المتغيرين . ولقد رأينا أيضا أنه إذا كان هناك متغيران مستقلان فإن التباين ، ومن ثم معامل الارتباط يكون صفرا . ودعنا الآن نثبت أن العكس ليس صحيحا ، أي أن متغيرين قد يكونا مرتبطين إرتباطا غير خطية ، وعلى الرغم من ذلك ، فإن قيمة معامل الارتباط بينهما قد تساوي صفرا . ونعيد التأكيد هنا على أن معامل الارتباط هو مقياس للعلاقة الخطية بين متغيرين .

جدول (١-٢)

X	$\rho(X)$
-1	$\frac{1}{3}$
0	$\frac{1}{3}$
1	$\frac{1}{3}$

فعلى سبيل المثال، افترض أن X متغير عشوائي وأن $Y = X^2$ ، عندئذ يصبح من الواضح أن X و Y مرتبطان ارتباطاً تاماً، حيث إن معرفة قيمة X تمكننا من التنبؤ تماماً بقيمة Y . وافترض الآن أن الدالة الاحتمالية للمتغير X موضحة بالجدول (١-٢)، أي أن X تأخذ القيم $-1, 0, +1$ بالاحتمال نفسه، دعنا الآن نحدد التغير بين X و Y :

$$\sigma_{X,Y} = E[(X - \mu_X)(Y - \mu_Y)]$$

ومن الجدول (١-٢)، نرى أن $\mu_X = 0$ ، ومن ثم، فإن المعادلة الخاصة بـ $\sigma_{X,Y}$ تصبح:

$$\sigma_{X,Y} = E[X(Y - \mu_Y)] = E(XY) - E(X\mu_Y) \quad (2.13)$$

والآن $Y = X^2$ افتراضاً، فإن:

$$\sigma_{X,Y} = E(X^3) - \mu_Y E(X) = E(X^3) \quad (2.14)$$

حيث $E(X) = 0$. ولايجاد $E(X^3)$ ، فنحن نلاحظ أن X^3 تأخذ القيم نفسها و احتمالات الحدوث الخاصة بـ X في الجدول رقم (١-٢). ومن ثم، يكون لدينا:

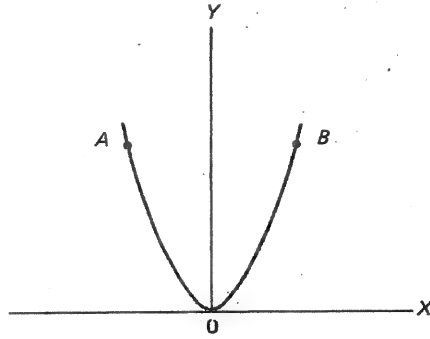
$$E(X^3) = -1\left(\frac{1}{3}\right) + 0\left(\frac{1}{3}\right) + 1\left(\frac{1}{3}\right) = 0 \quad (2.15)$$

ولذا، فإن $\sigma_{X,Y} = 0$ صفرًا، ويتبع هذا أن تكون $\rho_{X,Y}$ مساوية للصفر أيضاً.

ولنرى، بديهياً، ما يحدث دعنا نأخذ الحالة الأكثر عمومية حيث $Y = X^2$ ، ولكن X تأخذ كل القيم الممكنة بشرط أن تكون دالتها الاحتمالية متماثلة حول الصفر. ويعني الجزء الأخير من الجملة أن احتمال أن تقع قيمة X بين أي رقمين موجبين وليكونا 5، 10 هو نفسه احتمال أن تقع بين الرقمين السالبين المقابلين، -10، -5، ومع $Y = X^2$ يكون لدينا معادلة قطع مكافئ تتماس مع محور X عند نقطة الأصل كما هو موضح في شكل (٨-٢).

وقد يكون من الواضح الآن من الشكل (٨-٢) السبب الذي يجعل الارتباط بين متغيرين مثل X و Y صفرًا. بالطبع فإن كل المشاهدات عن X و Y تقع على القطع المكافئ لأن $Y = X^2$ ولأن الدالة الاحتمالية لـ X متماثلة حول الصفر، فإنه

لكل حدث مثل A يوجد هناك حدث مقابل له مثل B يمكن أن يقع بالاحتمال نفسه. ونتيجة لذلك، فإن الارتباط الطردي بين X وY عندما تأخذ X قيمة موجبة سوف يلغي الارتباط السلبي بينهما عندما تأخذ X قيمة سالبة. ولهذا، فإن الارتباط ككل بين X وY سوف يكون صفرا.



الشكل (٨-٢)

وفي مثل هذه الظروف توجد حالات بها ارتباط تام بين X وY، وعلى الرغم من ذلك فإن معامل الارتباط بينهما مساوي للصفر. وقبل ذلك، أوضحنا أن $\rho_{X,Y}=0$ عندما يكون X وY مستقلين. ونؤكد هنا أن $\rho_{X,Y}=0$ هو شرط ضروري وليس كافيا لأن يكون متغيران ما مستقلين. وبمعنى آخر إذا كانت $\rho_{X,Y}=0$ فإن X وY قد لا يكونان بالضرورة، مستقلين. ولكن لو أن X وY مستقلان فإن $\rho_{X,Y}=0$ بالضرورة. وكل ما يعنيه هذا من وجهة نظر التحليل القياسي (كما سوف نناقش فيما بعد) هو لو أن هناك ارتباطا ضعيفا بين متغيرين فلا بد أن نستبقي إمكانية وجود علاقة غير خطية بينهما.

مقدر معامل الارتباط

كما هو الأمر في حالة التغير، فإننا لانعرف بوجه عام معامل ارتباط المجتمع $\rho_{X,Y}$ وسوف تظهر لدينا مرة أخرى مشكلة تقدير. والمقدر البارز لـ $\rho_{X,Y}$ هو*:

* في بعض الكتب يستخدم الرمز r بدلا من الرمز ρ ، ولكننا نقضل استخدام $\hat{\rho}$ لتؤكد على أنه مقدر لـ ρ .

$$\hat{\rho}_{X,Y} = \frac{\hat{\sigma}_{X,Y}}{\hat{\sigma}_X \hat{\sigma}_Y} \quad (2.16)$$

حيث إن $\hat{\sigma}_Y$ و $\hat{\sigma}_X$ هما المقدران المعتادان للانحرافين المعياريين لكل من Y و X :

$$\hat{\sigma}_X = +\sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}} \quad \text{و} \quad \hat{\sigma}_Y = +\sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n-1}} \quad (2.17)$$

والآن دعنا نتعرف على خصائص $\hat{\rho}_{X,Y}$ في حالة العينة الكبيرة. فلقد لاحظنا سابقاً أنه كلما اقتربت n من مالانهاية كلما تقاربت \bar{Y} ، \bar{X} في الاحتمال μ_X و μ_Y على التوالي. وكنتيجة لذلك فإن $\sum (X_i - \bar{X})^2 / n-1$ تصبح هي متوسط الانحرافات المربعة للمتغير X عن وسطه الحسابي بناءً على بيانات عينة لانهاية الحجم. ومن ثم فإنها تتقارب في الاحتمال من تباين X والذي يرمز له بـ σ_X^2 . ويتبع هذا أن $\hat{\sigma}_Y$ تتقارب في الاحتمال من σ_X (وكذلك الأمر بالنسبة لـ $\hat{\sigma}_Y$). وطالما أن $\hat{\sigma}_{X,Y}$ تتقارب في الوقت نفسه من $\sigma_{X,Y}$ ، نرى في الأقل، بديهياً أن $\hat{\rho}_{X,Y}$ مقدر متسق لـ $\rho_{X,Y}$.

وعلى العكس من $\hat{\sigma}_{X,Y}$ ، فإن $\hat{\rho}_{X,Y}$ ليس مقدر غير متحيز عموماً. أي أن:

$$E(\hat{\rho}_{X,Y}) \neq \rho_{X,Y} \quad (2.18)$$

والسبب في هذا هو أن $\hat{\rho}_{X,Y}$ مشتق على أساس دالة غير خطية للمتغيرات $\hat{\sigma}_Y^2$ ، $\hat{\sigma}_X^2$ ، $\hat{\sigma}_{X,Y}$. وعلى وجه التحديد، فإن:

$$\hat{\rho}_{X,Y} = \frac{\hat{\rho}_{X,Y}}{\sqrt{\hat{\sigma}_X^2} \sqrt{\hat{\sigma}_Y^2}} \quad (2.19)$$

والآن يمكن إثبات أن $E(\hat{\sigma}_X^2) = \sigma_X^2$ و $E(\hat{\sigma}_Y^2) = \sigma_Y^2$. ولكن بتتبع مناقشتنا للدوال غير الخطية في الملحق B للفصل الأول، فإننا نجد أن*

* المناقشة التالية ليست إثباتاً رياضياً ولكنها توضح، ببساطة، أن $\hat{\rho}_{X,Y}$ متحيز.

$$E(\hat{\rho}_{X,Y}) = E\left(\frac{\hat{\sigma}_{X,Y}}{\sqrt{\hat{\sigma}_X^2} \sqrt{\hat{\sigma}_Y^2}}\right) \quad (2.20)$$

$$\neq \frac{E(\hat{\sigma}_{X,Y})}{\sqrt{E(\hat{\sigma}_X^2)} \sqrt{E(\hat{\sigma}_Y^2)}} = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y} = \rho_{X,Y}$$

وباختصار فإن $\hat{\rho}_{X,Y}$ مقدر متحيز لـ $\rho_{X,Y}$ ولكنه متسق. وهذا يعني أن التحيز يمكن اعتباره غير مهم إذا كان حجم العينة كبيرا، لأن خاصية الاتساق تؤكد لنا أن هناك احتمالا كبيرا بأن تكون $\hat{\rho}_{X,Y}$ قريبة من $\rho_{X,Y}$. وكما سوف نرى فيما بعد فإن مشكلة المقدرات المتحيزة والمتسقة سوف تظهر باستمرار في نماذج التقدير القياسية.

ملاحظة حول درجات الحرية

يتعين الإشارة إلى أن مقامي المقدرين $\hat{\sigma}_X^2$ ، $\hat{\sigma}_Y^2$ المعرفين بالمعادلة (2.17) يعرفان بدرجات الحرية، كما هو الحال في $\hat{\rho}_{X,Y}$ ، أي أنه، وكما في السابق، فعلى الرغم من وجود n في المجموعتين المقابلين لـ $\hat{\sigma}_X^2$ ، $\hat{\sigma}_Y^2$ ، فإنه يوجد $(n-1)$ معلومة مستقلة، فقط. وهذا يمكن رؤيته بديهيا للمتغير X بملاحظة مايلي:

$$\sum_{t=1}^n (X_t - \bar{X}) = 0$$

وعليه، فإن

$$(X_n - \bar{X}) = -(X_1 - \bar{X}) - \dots - (X_{n-1} - \bar{X})$$

وبمعنى آخر، فإن الحد الأخير من المجموع يعتمد، تماما، على الـ $(n-1)$ حد الأولى ومن، ثم فهو لا يحتوي على أية معلومات جديدة. ولقد تصادف أن تكون القسمة على درجات الحرية (وليس على عدد المشاهدات) هي الإجراء المعروف للحصول على مقدر غير متحيز لتباين متغير ما.

ولحسن الحظ، فإنه، لأغراض التعميم يمكن الحصول على درجات الحرية لمقدر التباين (من النوع المستخدم في هذا الكتاب) باستخدام قاعدة بسيطة. فعلى وجه التحديد، تساوي درجات الحرية لمثل هذا المقدر $(n-k)$ حيث n هي حجم العينة و k هي عدد المعلمات التي لابد من تقديرها لتحديد قيمة بسط المقدر. وعلى سبيل المثال، فإن مقدر التباين $\hat{\sigma}_x^2$ يوجد \bar{X} في بسط معادلته لأن μ_x غير معروفة. وفي هذه الحالة، $k=1$ ، ولو أن μ_x كانت معروفة لأمكن تقدير تباين X باستخدام المعادلة التالية:

$$\sum_{i=1}^n \frac{(X_i - \mu_x)^2}{n}$$

كلمة تحذير ^{٢٥} ص ١

قبل أن نبدأ في شرح نموذج الانحدار الخطي، فإن هناك كلمة تحذير يتعين قولها بشأن تفسير معامل الارتباط. ويلاحظ أننا كنا بصدد قياس درجة الاقتران الإحصائي بين متغيرين. ولم نقل شيئاً عن أي علاقة سببية بينهما. وفي حقيقة الأمر، من الممكن أن يوجد هناك ارتباط قوي بين متغيرين دون أن توجد أي علاقة سببية بينهما. فعلى سبيل المثال، من الممكن أن نجد ارتباطاً طردياً عبر الزمن بين متوسط المرتب السنوي للمدرسين بالدولار في الولايات المتحدة الأمريكية وبين الناتج الكلي للصلب. وهذا ربما لا يعكس أي نوع من التأثير المباشر لأحد المتغيرين على الآخر. ولكنه ببساطة راجع لحقيقة مؤداها أنه، لأسباب كثيرة، فإن كلا المتغيرين كانا يزدادان عبر الزمن في الحجم. ومن ثم فإن وجود ارتباط طردي أو عكسي (حتى ولو كان قوياً جداً) لا يعني أن هناك علاقة سببية بين المتغيرين.

مثال

على الرغم من أن وجود ارتباط قوي بين متغيرين لا يثبت أن هناك أي علاقة سببية بينهما، إلا أن هذا الارتباط قد يمدنا بتأييد عملي ذي قيمة لبعض

العلاقات المفترضة. فعلى سبيل المثال، دعنا نعود إلى ظاهرة من ظواهر التحضر التي يؤكد عديد من المراقبين أنها هي المصدر الرئيسي لمشاكل المدن المركزية. إنها ظاهرة نزوح الأسر متوسطة الدخل والمرتفعة من المدن إلى الضواحي*. والرأي هنا هو أن النزوح من المناطق الحضرية إلى الضواحي قد ترك مراكز المدن مأهولة بما تبقى من الأسر الفقيرة نسبياً، الأمر الذي أثبت أنه مكلف من وجهة النظر المالية، كما أنه ولد تدهوراً متزايداً لنوعية الحياة في المدن.

ولكن، هل الدليل الواقعي يؤيد هذا الرأي؟ ولإلقاء الضوء على هذا فإننا ربما نتساءل عما يمكن أن نتوقعه بشأن النتائج القابلة للملاحظة والقياس لهذه العملية، ثم نقوم بفحص البيانات الملائمة لتحديد ما إذا كانت، بالفعل، متسقة مع توقعاتنا. فعلى سبيل المثال، يمكننا مقارنة مدن عديدة بالولايات المتحدة لنرى ما إذا كانت هذه المدن التي قطعت فيها ظاهرة النزوح إلى الضواحي شوطاً طويلاً قد أصبحت في مركز سئ مقارنة بضواحيها. ويعرض جدول (٢-٢) بعض البيانات الخاصة بذلك حصل عليها من عينة مكونة من خمسة عشرة مدينة كبيرة بالولايات المتحدة. وبصورة محددة، دع P_i^c تشير إلى عدد سكان المدينة i ، ودع P_i^s تشير إلى عدد سكان منطقة الضواحي المحيطة بالمدينة رقم i . ويوضح الجدول (٢-٢) لكل واحدة من الخمس عشرة مدينة نسبة سكان منطقة الحضر التي تقيم في المدينة نفسها، أي $[P_i^c / (P_i^c + P_i^s)] \cdot 100$. ودع، أيضاً، Y_i^c و Y_i^s تشيران إلى متوسط دخل الأسرة في المدينة ومتوسط دخل الأسرة في ضواحي المدينة i على التوالي. ويعرض الجدول (٢-٢) بيانات عن نسبة الدخل بين المدينة والضواحي لكل واحدة من الخمس عشرة مدينة، أي $[Y_i^c / Y_i^s] \cdot 100$. ولو أن «فرضية النزوح» كانت صحيحة فإننا نتوقع أن المدن التي لديها نسبة ضئيلة نسبياً من سكان حاضرتها

* لدراسة قياسية مكثفة لهذه القضية، انظر:

David Bradford and Harry Kelejian. "An Econometric Model of the Flight to the Suburbs." *Journal of Political Economy*, 81 (May, June, 1973), pp. 566-589.

(أي قيمة منخفضة نسبياً لـ $100 [P_i^c / (P_i^c + P_i^s)]$ سوف تحتوي على أغلبية فقيرة دوماً في مركز المدينة. وبالنسبة لهذه المدن فإننا نتوقع أن تكون نسبة الدخل $100 (Y_i^c / Y_i^s)$ منخفضة نسبياً. وباختصار، فإننا نتوقع وجود اقتران بين المتغيرين $100 [P_i^c / (P_i^c + P_i^s)]$ و $100 (Y_i^c / Y_i^s)$.

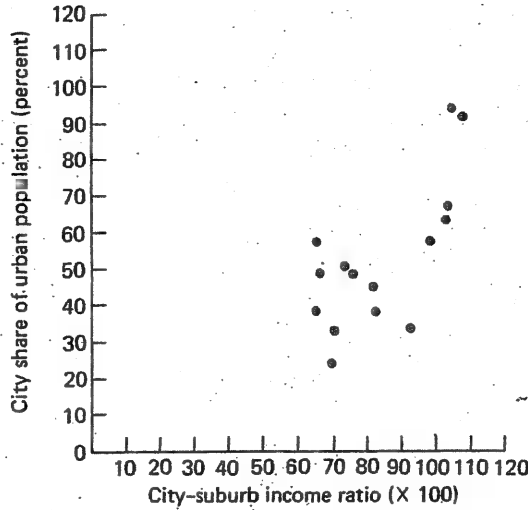
جدول (٢-٢)

المدينة	النصيب النسبي للمدينة من سكان الحضر % ^a	نسبة دخل المدينة إلى دخل الضواحي $\times 100$ ^b
بالتيمور	٥٧	٦٥
بوسطن	٢٤	٦٩
شيكاغو	٥٠	٧٣
كليفلاند	٣٨	٦٥
دالاس	٦٣	١٠٢
ديترويت	٣٨	٨٢
انديانابولس	٩١	١٠٧
لوس انجلوس	٣٤	٩٢
مفيس	٩٤	١٠٤
نيويورك	٤٩	٦٦
فيلادلفيا	٤٩	٧٥
فينكس	٦٧	١٠٣
سان دياغو	٥٨	٩٨
سانت لويس	٣٣	٧٠
سياتل	٤٥	٨١

^(a) عمود (٢) يشير إلى سكان مركز المدينة كنسبة من إجمالي سكان المنطقة الحضرية لعام ١٩٧٠.

^(b) عمود (٣) يوضح نسبة متوسط دخل الأسرة في مركز المدينة إلى متوسط دخل الأسرة في المنطقة خارج المدينة وداخل حدود المنطقة الحضرية لعام ١٩٦٩ (باستخدام تعريف التعداد المتفق عليه إحصائياً للمنطقة الحضرية).

والإجراء اللازم الآن هو أن نفحص البيانات لنرى ما إذا كانت توضح مثل هذا الاقتران الطردي أم لا، ومن الواضح أن شكل الانتشار (٢-٩) يوحي بذلك. فالقيم الأكبر لمتغير نسبة السكان تبدو في المتوسط مقترنة بالقيم الأكبر لمتغير الدخل النسبي، هذا على الرغم من أن نمط الاقتران ليس كاملاً.



شكل (٢-٩)

وعلى سبيل التوضيح للحسابات اللازمة، نقوم الآن بتقدير الارتباط بين متغيري نسبة السكان ونسبة الدخل. وللتبسيط، دع:

$$X_i = 100 \left[P_i^c / (P_i^c + P_i^s) \right] \text{ و } Y_i = 100 (Y_i^c / Y_i^s)$$

ومن ثم، فإنه، باستخدام (2.16)، نجد أن معامل ارتباط العينة بين هذين المتغيرين هو:

$$\begin{aligned} \hat{\rho}_{X,Y} &= \frac{\hat{\sigma}_{X,Y}}{\hat{\sigma}_X \hat{\sigma}_Y} = \frac{\Sigma (X_i - \bar{X})(Y_i - \bar{Y}) / n - 1}{\sqrt{\Sigma (X_i - \bar{X})^2 / n - 1} \sqrt{\Sigma (Y_i - \bar{Y})^2 / n - 1}} \\ &= \frac{\Sigma (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\Sigma (X_i - \bar{X})^2} \sqrt{\Sigma (Y_i - \bar{Y})^2}} \end{aligned} \quad (2.21)$$

ويمكن تبسيط الحسابات باستخدام بعض المتطابقات من ملحق A في الفصل الأول [على وجه التحديد (1A.10) و (1A.13)] لكي نضع (2.21) في صورة مختلفة لحد ما:

$$\hat{\rho}_{X,Y} = \frac{\sum X_i Y_i - n\bar{X}\bar{Y}}{\sqrt{\sum X_i^2 - n\bar{X}^2} \sqrt{\sum Y_i^2 - n\bar{Y}^2}} \quad (2.22)$$

وتظهر الحسابات الفعلية في جدول (٢-٣)، ويتضح في النهاية أن معامل ارتباط العينة بين المتغيرين هو 0.71. وبالتأكيد، فإن هذه النتيجة متسقة مع «فرضية النزوح».

جدول (٢-٣) حساب معامل ارتباط العينة

X_i	Y_i	$X_i Y_i$	X_i^2	Y_i^2	
57	65	3,705	3,249	4,225	
24	69	1,656	576	4,761	
50	73	3,650	2,500	5,329	
38	65	2,470	1,444	4,225	
63	102	6,426	3,969	10,404	
38	82	3,116	1,444	6,724	
91	107	9,737	8,281	11,449	
34	92	3,128	1,156	8,464	
94	104	9,776	8,836	10,816	
49	66	3,234	2,401	4,456	
49	75	3,675	2,401	5,625	
67	103	6,901	4,489	10,609	
58	98	5,684	3,364	9,604	
33	70	2,310	1,089	4,900	
45	81	3,645	2,205	6,651	
790	1,252	69,113	47,224	108,052	المجموع

$$\bar{X} = 52.7 \quad \bar{Y} = 83.5$$

$$\begin{aligned} \hat{\rho}_{X,Y} &= \frac{\sum X_i Y_i - n\bar{X}\bar{Y}}{\sqrt{\sum X_i^2 - n\bar{X}^2} \sqrt{\sum Y_i^2 - n\bar{Y}^2}} \\ &= \frac{69,114 - (15)(52.7)(83.5)}{\sqrt{47,224 - (15)(52.7)^2} \sqrt{108,052 - (15)(83.5)^2}} = 0.71 \end{aligned}$$

(٢-٢) وصف العلاقات السلوكية

لقد قدمنا في القسم السابق مقياسين للاقتتان الإحصائي بين متغيرين، ومع أخذهما في الحسبان نستمر الآن مع المسألة التي تعد ذات أهمية قصوى لنا: وهي

تحديد العلاقة الاقتصادية المفترضة وتقديرها. ولهذا الغرض نعود إلى مشكلة تقدير دالة الاستهلاك التي تعرضنا لها باختصار في الفصل الأول. فالنظرية الاقتصادية تفترض أن الإنفاق الاستهلاكي (C) دالة في الدخل المتاح (Y_d)، حيث إنه كلما زاد الدخل المتاح للأسرة زاد مستوى الإنفاق الاستهلاكي لها. ويفرض أن شكل هذه العلاقة الدالية خطي، كما يلي:

$$C_t = a + bY_{dt} \quad (2.23)$$

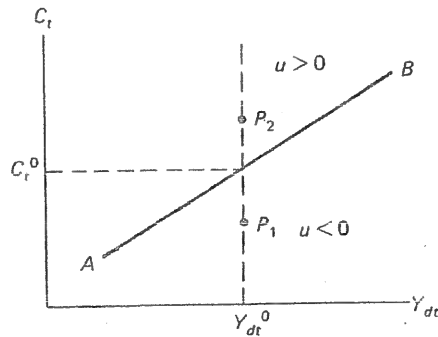
حيث (C_t) هي القيمة t للإنفاق الاستهلاكي و (Y_{dt}) هو القيمة t للدخل المتاح المقابل. فعلى سبيل المثال قد تشير t إلى الفترات الزمنية، وفي هذه الحال فإن المعادلة (2.34) تربط بين الإنفاق الاستهلاكي في الفترة الزمنية t ومستوى الدخل المتاح في الفترة نفسها. وكما قد تشير t إلى أفراد عند نقطة زمنية معينة. وفي هذه الحال فإن (2.23) تربط بين الإنفاق الاستهلاكي للفرد t ومستوى دخله المتاح.

ومن الجدير بالملاحظة أولاً أننا نحدد هنا علاقة سببية مفترضة. وتقول النظرية إن مستوى الإنفاق الاستهلاكي يعتمد على مستوى الدخل المتاح. فكلما زاد الدخل المتاح للأسرة، فمن المتوقع أن تنفق جزءاً من هذه الزيادة على الاستهلاك. وثانياً، فإن المعادلة (2.23) تحدد علاقة تامة بين (C_t) و (Y_{dt}). فلو أنه، على سبيل المثال، إذا كانت $a=100$ و $b=0.9$ فإن المعادلة (2.23) تقرر أن $C_t = \$13600$ لو أن $Y_{dt} = \$15000$. وعلى الرغم من ذلك، إذا نظرنا إلى البيانات بإمعان لن نجد علاقة تامة. فلا يوجد هناك مجموعة من النقاط التي تقع على طول خط مستقيم بدقة، ولكن، بدلاً من ذلك نجد انتشاراً من النقاط. فالعلاقة بين الاستهلاك والدخل المتاح التي نريد تقديرها ليست علاقة تامة، وإنما هي علاقة نمطية. ومانفكر فيه هو النوع غير المحدد مثل: «إذا كان الدخل المتاح للأسرة يساوي \$15000، فإنه في المتوسط سوف يكون الإنفاق الاستهلاكي للأسرة مساوياً لـ \$13600. وفي أي حالة خاصة، فنحن نتوقع لعديد من الأسباب التي سوف نناقشها بعد قليل - أن تنحرف C_t إما لأعلى

أو لأسفل عن القيمة النمطية. وهذا يتضمن أن دالة الاستهلاك البسيطة يمكن كتابتها كمايلي:

$$C_t = a + bY_{dt} + u_t \quad (2.24)$$

حيث u_t قد تأخذ قيما موجبة أو سالبة، ولكن، بمتوسط يساوي الصفر، حيث إن القيمة المتوسطة لـ C_t المقابلة لقيمة محددة لـ Y_{dt} تكون مساوية لـ $(a + bY_{dt})$. افترض، على سبيل المثال أن العلاقة المتوسطة بين C_t و Y_{dt} ممثلة بالخط AB في الشكل (١٠-٢). ومن ثم، لو أن $Y_{dt} = Y_{dt}^0$ ، فإن متوسط C_t سوف يكون C_t^0 . وعلى الرغم من ذلك فإنه عموما عندما $Y_{dt} = Y_{dt}^0$ فإن قيمة C_t سوف تنحرف بمقدار ما عن C_t^0 وفي بعض الحالات C_t و C_t^0 ، كما عند النقطة P_1 مما يتضمن أن $u_t < 0$ ، وفي حالات أخرى، تكون $C_t > C_t^0$ مع $u_t > 0$ كما هو عند النقطة P_2 في الشكل (١٠-٢).



شكل رقم (١٠-٢)

والحد u_t الذي هو ذو أهمية كبرى في الاقتصاد القياسي يسمى حد الخطأ disturbance (or error) term (أو أكثر شيوعاً الخطأ العشوائي). إنه الوسيلة التي تمكننا من توضيح أن العلاقات الاقتصادية ليست تامة (مؤكدة) ولكنها تمثل أنماطا سلوكية

متوسطة. ويشير هذا قضية فحواها: لماذا لانجد في الاقتصاد علاقات دقيقة تنطبق بدون استثناءات.* أي لماذا تختلف u دائما عن الصفر؟ ويرجع هذا لعدد من الأسباب أهمها:**

(١) المتغيرات المحذوفة

دعنا نعود مرة أخرى إلى حالة الإنفاقات الاستهلاكية، فلاشك أن الاستهلاك يعتمد على عدد من المتغيرات الأخرى غير الدخل المتاح. وعلى سبيل المثال إذا كانت C_t تشير إلى الإنفاقات الاستهلاكية في الفترة t ، فإننا قد يمكننا افتراض أن:

$$C_t = a + b_1 Y_{dt} + b_2 L_t + b_3 \dot{P}_t + b_4 R_t + \dots, \quad (2.25)$$

حيث L_t تشير إلى رصيد الأصول السائلة في الفترة t ، \dot{P}_t تشير إلى التغير النسبي في الأسعار خلال الفترة t و R_t تشير إلى معدل الفائدة خلال الفترة t وهكذا. ومن ناحية أخرى، فإننا قد نشعر أن Y_{dt} هو، في الأقل، أهم العوامل المحددة للاستهلاك C_t تحت الظروف العادية، وأن آثار المتغيرات الأخرى سوف تكون ضعيفة وسوف يلغى بعضها بعضا مع مرور الزمن.

وفي هذه الحالة، فإن الخطأ العشوائي يمثل مجموع كل هذه الحدود المحذوفة:***

$$u_t = (b_2 L_t + b_3 \dot{P}_t + b_4 R_t + \dots) \quad (2.26)$$

* ينبغي علينا الإشارة إلى وجود بعض العلاقات التامة في الاقتصاد، ولكنها ليست علاقات سلوكية إقترحتها النظرية الاقتصادية، وتعرف هذه «بالتطابقات المحاسبية»، وتعتبر صحيحة بالتعريف. فعلى سبيل المثال، فإن واحدة من التطابقات المحاسبية المعروفة في الاقتصاد هي تقرير الميزانية الأساسي والذي يعتبر أن:

$$\text{الأصول} = \text{الخصوم} + \text{الرصيد الصافي}$$

وتعد هذه العلاقة صحيحة دائما بسبب الطريقة التي يعرف بها الرصيد الصافي:

$$\text{الرصيد الصافي} = \text{الأصول} - \text{الخصوم}$$

فالرصيد الصافي هو المقدار المتبقي الذي يؤكد أن الميزانية متوازنة.

** تعتمد المناقشة التالية على:

J. Johnston. *Econometric Methods*, 2nd ed. New York: McGraw. Hill, 1972, pp. 10-11.

*** سوف نتعرض لاحقا للحالة التي لا تميل فيها هذه المتغيرات المحذوفة لإلغاء أثر بعضها بعضا.

وباختصار فإن حد الخطأ قد يظهر لأن كل العوامل المؤثرة قد لا يمكن أخذها جميعاً في الاعتبار.

(٢) السلوك غير المتوقع للأفراد

إن الأنماط السلوكية للناس نادراً ما يمكن التنبؤ بها على وجه الدقة. ومن ناحية أخرى، فإن السلوك البشري ليس ذا صفة عشوائية بحثة عموماً. وفي هذا الصدد يمكننا أن ننظر إلى نموذج الاستهلاك (2.24) على أنه يتضمن جزئين: جزءاً محدداً وهو الذي يربط الإنفاق بالدخل $(a + bY_{dt})$ وجزءاً لا يمكن التنبؤ به (أو غير محدد) u . وفي هذا الإطار، فإن حد الخطأ يعكس أو يأخذ في الاعتبار «الاحتياجات الطارئة»، و «التغيرات في الرأي»، أو التغيرات في الموقف التي تحفز المستهلكين على أن ينفقوا أكثر أو أقل مما اعتادوا عليه. وهذا المقدار الذي اعتادوا على إنفاقه يمثل في الجزء المحدد $(a + bY_{dt})$. وفي الإطار نفسه، فإن حد الخطأ يمكن النظر إليه على أنه يعكس آثار الأحداث التي لا يمكن التنبؤ بها على السلوك الاقتصادي. فعلى سبيل المثال، لو أن صديقاً من خارج المدينة قام بزيارة غير متوقعة، فإن مضيفه قد يتجاوز ميزانيته العادية بأن يأخذ صديقه لعشاء في الخارج.

(٣) تباين سلوك الأفراد

وبالمثل نعرف، جميعاً، أنه بسبب الاتجاهات المختلفة، يكون لبعض الأسر ميل أكبر للادخار من الأسر الأخرى. ولو أننا استخدمنا المعادلة (2.24) لنفسر الإنفاقات المختلفة للأسر عند نقطة زمنية معينة يمكننا، اعتبار $(a + bY_{dt})$ هي إنفاق الأسرة العادية (المتوسطة) التي دخلها هو Y_{dt} ونعتبر حد الخطأ u_t هو الانحراف عن المتوسط. أي أن u سوف تعكس المواقف المختلفة تجاه الادخار: فالأسر ذات الميل المرتفع للادخار نسبياً تقابل القيم السالبة لـ u_t ، والأسر ذات الميل المنخفض للادخار سوف تقترن بالقيم الموجبة لـ u_t .

(٤) أخطاء القياس

حتى لو كانت $C_t = a + bY_{dt}$ مؤكدة، فإنه قد لا يمكننا قياس C_t بدقة كاملة. ونتيجة لمثل هذه الأخطاء في القياس، فقد نشاهد في الواقع \bar{C}_t المرتبطة بـ C_t كما تصفها الصيغة التالية:

$$\bar{C}_t = C_t + u_t$$

وتمثل u_t هنا خطأ القياس. وهذا يعني أنه، بينما نقلل من قيمة C_t أو نبالغ فيها تبعاً لما إذا كانت u_t موجبة أو سالبة، فإننا نفترض أن أخطاء القياس تميل إلى أن يلغى بعضها بعضاً، فلو أخذنا قياسات متكررة لـ C_t فإن متوسط هذه القياسات من المتوقع أن يكون C_t . ولو احللنا $\bar{C}_t = C_t + u_t$ في $C_t = a + bY_{dt}$ نحصل على:

$$\bar{C}_t = a + bY_{dt} + u_t$$

وبمعنى آخر حتى لو كانت العلاقة بين الإنفاق الاستهلاكي والدخل المتاح مؤكدة، فإن العلاقة بين الإنفاق المقاس والدخل قد تتضمن خطأ عشوائياً.*

ولهذا، فإننا سوف نكون العلاقات الدالية التي تصف السلوك الاقتصادي بأنه علاقات متوسطة، وسوف نوضح هذا بإدخال خطأ عشوائي، u_t في النموذج. ومن الأمثلة الأخرى على تلك العلاقات الاقتصادية المثالان التاليان:

$$I_t = e + fR_t + u_t \quad (2.27)$$

$$Q_t = g + hL_t + u_t \quad (2.28)$$

حيث I_t = الاستثمار، R_t = سعر الفائدة، Q_t = مستوى الناتج و L_t = عنصر العمل مقاساً بساعات عمل. ويتضح أنه، في كل واحدة من هذه المعادلات أن هناك متغيرات إضافية أخرى مهمة تؤثر في المتغير التابع: فمن الواضح أن الاستثمار يعتمد على متغيرات أخرى بجانب مستوى معدلات الفائدة كالطلب على الناتج. وفي الحالة الثانية فإن مستوى الناتج يتغير مع تغير كميات المدخلات الأخرى التي

* لاحظ أننا، لأغراض التبسيط افترضنا عدم وجود أخطاء في قياس المتغير المستقل.

تستخدم مع العمل في التوليفة نفسها. ولهذا السبب وحده، نتوقع عدم دقة العلاقات بين هذه المتغيرات.

(٢-٣) نموذج انحدار المتغيرين

افترض أننا كونا علاقة سلوكية من النوع الذي سبقت مناقشته، وبشكل أعم افترض أن لدينا العلاقة الخطية التالية:

$$Y_t = a + bX_t + u_t, \quad t = 1, \dots, n \quad (2.29)$$

حيث:

Y_t المشاهدة رقم t للمتغير التابع،

X_t المشاهدة رقم t للمتغير المستقل،

u_t القيمة المقابلة للخطأ العشوائي t و a ، b هما معلمتان مجهولتا القيمة.

والمعادلة (2.29) علاقة خطية بين Y_t ، X_t و u_t بها معلمتان مجهولتان هما a و b ونفترض أن هذه العلاقة تتحقق لجميع القيم المحددة عند t أي $t=1,2,\dots,n$. لاحظ أن قيم t تقابل الـ n مشاهدة لكل من Y_t و X_t . ومن المتوقع أن تتغير قيمة الخطأ العشوائي u_t في العلاقة السابقة من مشاهدة لأخرى لتعكس الانحرافات عن أنماط السلوك النمطية. وعلى العكس من قيم Y_t و X_t ، فنحن لانفترض أن قيم الخطأ العشوائي قابلة للملاحظة.

وتتمثل المشكلة الأولى في تقدير قيمتي المعلمتين a و b حتى تتمكن من عمل تقدير كمي للعلاقة بين Y_t و X_t . ولإتمام ذلك لابد أولاً من وضع عدد من الافتراضات الرياضية الخاصة بالطريقة التي نحصل بمقتضاها على كل من قيم المتغير المستقل X_t وقيم الخطأ العشوائي u_t . وسوف نتضح فيما بعد ضرورة هذه الافتراضات خاصة تلك المتعلقة بـ u_t . فعلى سبيل المثال، طالما أن Y_t تعتمد على كل من X_t و u_t ، فإنه يتبع ذلك أن تعتمد طبيعة أي علاقة بين Y_t و X_t على خصائص الخطأ العشوائي u_t .

الافتراضات الأساسية The Basic assumptions

١- نفترض أولاً أن قيم المتغير المستقل X ليست متساوية. ففي الأقل، لا بد أن توجد هناك قيمة واحدة تختلف عن البقية. وكما سنرى، فمالم يتحقق هذا الشرط فلن نستطيع تقدير a و b . ولحد ما، فإنه من البديهي إذا لم تتغير X مطلقاً فإننا لن نتمكن من مشاهدة الكيفية التي تتغير بها Y مع X .

ويشير هذا تساؤلاً عن ما الذي يحدد القيم الخاصة بـ X . والافتراض التقليدي هو أن القائم بالتجربة نفسه يختار قيم X وعندئذ، يشاهد القيم المقابلة أو الناتجة لـ Y . فعلى سبيل المثال، دع المعادلة (2.29) تمثل العلاقة بين الناتج من القيم بالبوشل bushel لكل فدان أرض (Y)، وعنصر السماد للفدان (X) مقاساً بالرطل. ولاستكشاف هذه العلاقة، فإن القائم بالتجربة قد يضع $X=1$ في الفدان الأول من الأرض و $X=2$ في الفدان الثاني (أي $X_1=1, X_2=2$) وهكذا.

غير أنه من الثابت أن الاقتصاديين ليسوا، عادة، محظوظين بهذا الشكل. افترض، على سبيل المثال، أننا نريد أن نفحص العلاقة بين معدل التضخم ومعدل البطالة. في هذه الحالة، يمكننا تعريف Y في المعادلة (2.29) بأنها التغير النسبي في الأسعار من عام لآخر و X بأنها التغير النسبي في قوة العمل العاطلة خلال الفترة نفسها. ومن الواضح أنه لا يمكننا الاستمرار في تحرى العلاقة بين X و Y عن طريق وضع معدل بطالة عند نسبة مختلفة كل عام ثم ملاحظة معدل التضخم الناتج عن ذلك. فمعدل البطالة يتحدد بحركة الاقتصاد ككل. ولهذا السبب، فهو خارج تحكمنا التجريبي، حيث لا يمكننا اختيار قيم المتغير المستقل وتغييره بانتظام. ومن ثم، فلا بد أن نشاهد في هذه الحالة كلا من X و Y معاً. ولذا، فسوف نواصل مناقشتنا مع افتراض أنه، مهما كانت الآلية التي تنتج عنها قيم X ، فإنها سوف تعطي في الأقل، قيمتين مختلفتين للمتغير X .

٢- الافتراضات المتعلقة بخصائص الخطأ العشوائي نفسه:

$$2a. E(u_t) = \mu_u = 0$$

$$2b. E(u_t - \mu_u)^2 = E(u_t^2) = \sigma_u^2,$$

2c. u_t is independent of u_s for $s \neq t$, and so

$$E[(u_t - \mu_u)(u_s - \mu_u)] = \text{COV}(u_t, u_s) = 0$$

ويقرر الافتراض 2a أنه، لكل مشاهدة، فإن القيمة المتوقعة للخطأ العشوائي تساوي الصفر. وعلى سبيل مثال بسيط، نفترض أنه لكل مشاهدة، تتحدد قيمة u كمايلي: شخص غير معروف لنا رمي عملة، فلو ظهرت الشارة فإنه يضع $u_t=1$ ، ولو ظهرت الكتابة فإن يضع $u_t = -1$. وفي هذه الحالة:

$$E(u_t) = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$$

والمنطق وراء افتراض أن متوسط الخطأ العشوائي يساوي الصفر منطق مباشر. فنحن نفترض أن نظريتنا المتمثلة في المعادلة (2.29) تصف، بدقة، السلوك المتوسط للمتغير Y والمقابل للقيم المختلفة لـ X . أي أنه مهما كانت قيمة X ، فإن القيمة المتوسطة لـ Y سوف تكون $Y^m = (a + bX)$. وعلى العكس من ذلك لو أن $E(u_t) \neq 0$ فإن الأمر لن يكون كذلك. أفترض، مثلاً، أن u_t في المعادلة (2.29) كانت دائماً موجبة، ومن ثم، فإن هذا سوف يتضمن أن القيمة المتوسطة لـ Y المقابلة لكل قيمة من قيم X سوف تفوق $(a+bX)$ لذلك، فإن المعادلة (2.29) يمكن النظر إليها على أنها مشتقة من معادلة أخرى متوسط الخطأ العشوائي فيها لايساوي صفر. فعلى سبيل المثال، افترض أن $E(u_t) = d \neq 0$ حيث d ثابت، عندئذ، يمكن تعريف:

$$v_t = u_t - d, \quad (2.30)$$

ولاحظ أن $E(v_t) = E(u_t) - d = d - d = 0$. وباستخدام (2.30) فإننا نستطيع إحلال $u_t = v_t + d$ في (2.29) لنحصل على:

$$\begin{aligned} Y_t &= (a + b) + bX_t + v_t \\ &= a^* + bX_t + v_t, \end{aligned} \quad (2.31)$$

حيث $a^* = a + d$ ، وكما هو ملاحظ فإن $E(v_t) = 0$. ولذا فإننا يمكن أن نأخذ (2.30)، على أنه نموذجنا للانحدار.

ويوضح الافتراض (2b) أن تباين الخطأ العشوائي ثابت ويساوي σ_u^2 . ولذا فهو لا يتغير بانتظام مع تغير t . فلو استخدمنا سلسلة من المشاهدات عن الإنفاقات الاستهلاكية والدخل المتاح عبر الزمن، فإن هذا الافتراض يعني أن تباين الحد العشوائي لا يزيد ولا ينقص مع مرور الزمن. * أو باستخدام مثالنا البسيط السابق، فإن هذا الافتراض سوف يخرق لو أنه عند رمي قطعة العملة، قام الشخص غير المعروف لنا بوضع $u_t = +t$ ووضع $u_t = -t$ عندما يظهر الشعار أو الكتابة على التوالي. وفي هذه الحالة، سوف تظل القيمة المتوقعة لـ u مساوية للصفر.

$$E(u_t) = \frac{1}{2}(t) + \frac{1}{2}(-t) = 0$$

ولكن، من الواضح أن تباين u_t سوف يصبح أكبر مع كل مشاهدة تالية. أما المنطق وراء هذا الافتراض والطرق التي تعالج مشاكل التقدير التي تظهر عندما يخرق هذا الافتراض، فسوف نتعرض لها رياضيا فيما بعد بالكتاب (انظر الفصل السادس). وعند هذه النقطة، فإننا نلاحظ أنه لو لم يكن تباين u_t متساويا عند جميع المشاهدات، فإن جميع المشاهدات لا يمكن الاعتماد عليها أو الوثوق بها بالدرجة نفسها. فعلى سبيل المثال، افترض أننا عرفنا أن تباين u_t كان مساويا الصفر لمشاهدين معينتين ولتكونا الأولى والثانية، ولكنه كان موجبا للمشاهدات الأخرى. عندئذ، فطالما أن متوسط u_t يساوي الصفر، فإن هاتين المشاهدين مع كون احتمالهما مساويين الواحد الصحيح يحققان المعادلة لتالية $Y = a + bX$ ، ويمكننا كتابة:

$$\begin{aligned} Y_1 &= a + bX_1, \\ Y_2 &= a + bX_2, \end{aligned} \quad (2.32)$$

* ولو كنا، بدلا من ذلك، نبحث مسح ميزانية يوضح الإنفاقات الاستهلاكية للأسر ذات الدخل المختلفة، فإن هذا الافتراض سوف يتضمن أن تباين حد الخطأ σ_u^2 لا يتغير بانتظام مع تغير الدخل المتاح. انظر الفصل السادس.

وعندئذ، نحل المعادلتين في (2.32) لـ a و b . وبمعنى آخر يمكننا أن نهمل كل المشاهدات الأخرى ونقدر a و b ببساطة باستخدام هاتين النقطتين الأوليين. وباختصار، فإن المشاهدات التي تقابل تباينات صغيرة ذات أهمية أكبر لحد ما من المشاهدات التي تقابل تباينات كبيرة. ومن أجل ذلك تكون جميع مشاهداتنا على القدر نفسه من الأهمية في هذه المرحلة فلقد وضعنا الافتراض (2b).

ويقرر الافتراض (2c) أن قيمة الخطأ العشوائي رقم t مستقلة عن قيمة أي خطأ عشوائي آخر وليكن رقم s . والسبب وراء هذا الافتراض هو أننا، على وجه التحديد، نريد أن نعين نموذج توجد فيه قوة منتظمة واحدة متنبأ بها قابلة للتنبؤ (X_t) تؤثر على المتغير التابع Y_t . ولو أن الأخطاء العشوائية كانت مرتبطة مع بعضها البعض فإنه من الواضح أن هذا لن يحدث. فعلى سبيل المثال، افترض أن u_t كانت مرتبطة عكسيا مع القيمة السابقة لها مباشرة u_{t-1} . عندئذ، فإن قيمة Y_t سوف تعتمد بانتظام ومتنبأ بها على قيمة X_t وقيمة u_{t-1} طالما أن u_{t-1} تحدد u_t ، في الأقل جزئيا. وعلى الرغم من أننا سوف نتناول مثل هذه النماذج في الفصل السادس فسوف نبدأ منقشتنا لتحليل الانحدار على مستوى أبسط بافتراض أن قيمة الخطأ العشوائي لأي مشاهدة لا تعتمد على قيمته عند أي مشاهدة أخرى.

ومع هذه المجموعة من الافتراضات، نكون قد وصفنا حد الخطأ في المعادلة (2.29) بأنه متغير عشوائي غير قابل للمشاهدة، وسطه الحسابي يساوي صفراً، وتباينه ثابت σ_u^2 بالإضافة إلى أن قيمه في أي ظرف معين مستقلة. ولذا غير مرتبطة مع قيمه في أي ظروف أخرى.

٣- وافترضنا النهائي هو أن u_t مستقلة عن كل القيم n للمتغير المستقل X . ويتبع هذا أن $\text{cov}(u_t, X_t) = 0$ ، وحيث إن $E(u_t) = 0$ من الافتراض 2a، فإننا نحصل على:

$$\begin{aligned}\text{cov}(u_t, X_t) &= E[(u_t - 0)(X_t - \mu_X)] = E(u_t X_t) - E(u_t \mu_X) \\ &= E(u_t X_t) - \mu_X E(u_t) = E(u_t X_t) = 0\end{aligned}$$

ومن ثم، فإن الافتراضات تتضمن أن $E(u_t, X_t) = 0$.

والمنطق وراء الافتراض بأن X_i و u_i مستقلان*، ومن ثم، $\text{cov}(u_i, X_i) = 0$ ، مشابهة للمنطق نفسه وراء الافتراض $E(u_i) = 0$ فلو أن أي واحد من هذين الافتراضين لم يتحقق فإن القيمة المتوسطة للمتغير Y_i^m ، مثل Y_i^m والمقابلة لقيمة معينة للمتغير X_i لن تكون، عموماً، هي $(a + bX_i)$. دعنا نأخذ على سبيل المثال الحال التي يرتبط فيها X_i و u_i طردياً. ومن ثم، فإن هذا الارتباط الطردي يتضمن أن قيم u_i الأكبر من المتوسط [الموجبة حيث إن $E(u_i) = 0$] سوف تميل للاقتزان بقيم X_i الأكبر من المتوسط، وبالطريقة نفسها سوف تميل قيم u_i الأقل من المتوسط (السالبة) للاقتزان بقيم X_i الأقل من المتوسط. ويتضمن هذا أن متوسط u_i المقابل لقيم X_i الكبيرة فقط، سوف يكون موجبا، وعلى العكس من ذلك، فإن متوسط u_i المقابل لقيم X_i الصغيرة فقط سوف يكون سالبا. ومن هذا فإننا نرى أن متوسط Y_i أي Y_i^m يفوق $(a + bX_i)$ عندما تكون X_i صغيرة.

ونوضح هذه المشكلة في الشكل (٢-١١). افترض أن الخط AB يمثل العلاقة $Y_i^* = a + bX_i$ ، وبنفس الطريقة دع الخط CD يشير إلى متوسط قيمة Y_i (الذي يرمز لها Y_i^m) والتي تقابل القيم المختلفة لـ X_i . وعندئذ، فإن الارتباط الطردي بين الخطأ العشوائي والمتغير المستقل يتضمن أن العلاقة بين Y_i^m و Y_i^* سوف تكون، إلى حد ما، كما هي موضحة بالشكل (٢-١١). **

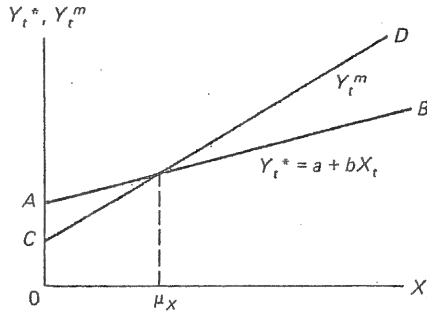
وبهذا نستكمل مناقشتنا حول الافتراضات، فتعين العلاقة بين Y_i و X_i في الصيغة التالية:

$$Y_i = a + bX_i + u_i$$

بالإضافة إلى الافتراضات التي ناقشناها، تكون نموذجنا الخطي للانحدار. ومهمتنا التالية هي أن نرى كيف يمكننا استخدام افتراضاتنا للحصول على مقدرات لـ a و b . وفي مجرى مناقشتنا، فسوف نرى، بوضوح، ما الدور الدقيق لكل افتراض وكيف تعتمد نتائجنا عليها.

*أهمية الافتراض بأن u_i مستقل عن جميع الـ n قيمة للمتغير X هي أهمية فنية، وسوف نبينها فيما بعد.

** نحن نتجاوز هنا قليلاً لأنه في النظرية CD ليس من المتعين أن يكون خطأ مستقيماً.



شكل (١١-٢)

(٢-٤) تقدير معادلة الانحدار - طريقة المتغير المساعد

عرف المدخل الذي نستخدمه في الأدب الاقتصادي بتقدير المتغير المساعد وهي طريقة تتضمن تطبيق افتراضاتنا المتعلقة بنموذج الانحدار الأساسي مباشرة على القيم المشاهدة للعينة Y و X .^{*} وكما سوف نرى بعد قليل، فإن هذا يمكننا من توليد مقدرات لكل من a, b . وسر الإعجاب بهذا المدخل هو أنه يسمح لنا بأن نرى، بوضوح، الأهمية الخاصة أو الدور الذي يقوم به كل افتراض وضعناه في نموذج الانحدار.

دعنا نعود إلى معادلة الانحدار (2.29) الأساسية:

$$Y_t = a + bX_t + u_t, \quad t = 1, \dots, n$$

وبسبب افتراضنا أن $E(u_t) = 0$ أن القيمة المتوسطة لـ Y_t والمقابلة لقيمة معينة لـ X_t تصبح:

$$Y_t^m = a + bX_t \quad (2.33)$$

^{*} الصيغة الخاصة بطريقة المتغير المساعد التي سوف نستخدمها كان قدمها من قبل Arthur S. Goldberger في كتابه

والمعادلة (2.33) قد يمكن تفسيرها بوصفها متوسطة بين Y_t و X_t . ويظهر من (2.29) و (2.33) أن:

$$Y_t = Y_t^m + u_t \quad (2.34)$$

وتقرر المعادلة (2.34)، ببساطة، أن Y_t يمكن التعبير عنها باعتبارها مجموعاً لمكونين: مكون المتوسط، ومكون الحد الذي يسبب انحرافها عن المتوسط. وإعادة تنظيم حدود (2.34)، يمكن التعبير عن حد الخطأ كمايلي:

$$u_t = Y_t - Y_t^m \quad (2.35)$$

افترض الآن أن لدينا مقدرين لـ a و b وليكونا \hat{a} و \hat{b} . وفي ضوء (2.33)، فإن مقدر القيمة المتوسطة لـ Y_t سوف يكون:

$$\hat{Y}_t = \hat{a} + bX_t \quad (2.36)$$

حيث إننا قد بسطنا الصيغة بعدم الإشارة إلى الدليل العلوي من المقدر الخاص بـ Y_t^m . بالمثل، فإن مقدرنا للخطأ العشوائي الموضح بالمعادلة (2.35) سوف يكون:

$$\hat{u}_t = Y_t - \hat{Y}_t \quad (2.37)$$

أي من الممكن الحصول على مقدر للخطأ العشوائي من المعادلة (2.35) عن طريق إحلال المعلمتين المجهولتين a و b بمقدريهما. ومرة أخرى بإعادة ترتيب حدود (2.37) نحصل على معادلة مقابلة للمعادلة (2.34):

$$\begin{aligned} Y_t &= \hat{Y}_t + \hat{u}_t \\ &= \hat{a} + \hat{b}X_t + \hat{u}_t \end{aligned} \quad (2.38)$$

ويلاحظ أن (2.38) تعبر عن قيمة Y_t بدلالة مقدراتنا لـ \hat{a} ، \hat{b} و \hat{u}_t (أي a و b و u_t) وقيمة X_t .

دعنا الآن نعود إلى مشكلة الحصول على \hat{a} ، \hat{b} . فمن بين افتراضات نموذج الانحدار أن u_t لها متوسط يساوي الصفر $E(u_t) = 0$. وهذا يقودنا بدهيا إلى توقع أنه لو أمكننا الحصول على متوسط قيم u_t ذات العدد n أي $\bar{u} = \sum u_t / n$ فإن هذا المتوسط سوف يأخذ قيمة صغيرة. واصطلاحاً يمكننا القول إن $E(u_t) = 0$ يتضمن

أن $E(\bar{u}) = 0$. ويتضمن هذا كله أنه إذا كانت \hat{u}_t معرفة بالمعادلة (2.37) فإنه من المرغوب فيه أن يكون:

$$\left(\sum_{t=1}^n \frac{\hat{u}_t}{n} \right) = 0 \quad (2.39)$$

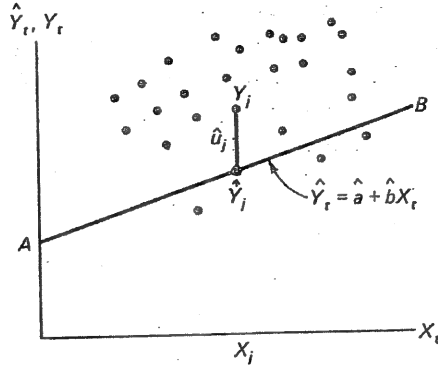
وبالضرب في n :

$$\sum_{t=1}^n \hat{u}_t = 0 \quad (2.40)$$

أي أننا قد نرغب في أن يتصف \hat{u}_t بخاصية (2.39) أو (2.40) والتي تقابل أحد افتراضاتنا الأساسية الخاصة بالخطأ العشوائي، أي $E(u_t) = 0$.
 \hat{u}_t وقبل الاستطراد قد يكون من المفيد تفسير معنى (2.40) جبرياً . فمن (2.38) نرى أنه إذا كان $\Sigma \hat{u}_t \neq 0$ فإن $\Sigma Y_t \neq \Sigma \hat{Y}_t$. ولغرض التوضيح افترض أن $\Sigma \hat{u}_t = 500$ ، ومن ثم، سوف يتبع ذلك أن تكون $\Sigma Y_t > \Sigma \hat{Y}_t$. والآن تمنع في الشكل رقم (2.12) والذي تمثل فيه الـ n مشاهدة لكل من Y_t و X_t بنقاط الانتشار . وتمثل المعادلة المقدرة بين Y_t و X_t أي:

$$\hat{Y}_t = \hat{a} + \hat{b}X_t$$

بالخط AB . ويلاحظ عموماً أن النقاط تقع فوق الخط . والسبب في ذلك هو أن ارتفاع الخط المقابل لقيمة معينة للمتغير المستقل ولتكن X_j هو \hat{Y}_j . ولكن عموماً سوف تكون هذه القيمة \hat{Y}_j أقل من القيمة المشاهدة للمتغير التابع Y_j طالما $\Sigma Y_t > \Sigma \hat{Y}_t$. ولقد أصبح من الواضح أنه إذا كانت $\Sigma \hat{u}_t$ سالبة فإنه بنفس المنطق سوف يقع انتشار النقاط عموماً تحت خط العلاقة المقدرة بين Y_t و X_t . ومن ثم، فإن الشرط (2.40) بأن $\Sigma \hat{u}_t = 0$ يتضمن أنه في المتوسط لا تتركز النقاط فوق الخط المقدّر أو أسفله .



شكل (٢-١٢)

وقد يكون من الواضح الآن ما دور (2.40) في الحصول على المقدرات a و b فلو جمعنا (2.38) عبر n مشاهدة، فسوف نحصل على:

$$\begin{aligned}\Sigma Y_t &= \Sigma \hat{Y}_t + \Sigma \hat{u}_t \\ &= n\hat{a} + \hat{b}\Sigma X_t\end{aligned}\quad (2.41)$$

طالما أن $\Sigma \hat{u}_t = 0$ من (2.40)، وبقسمة (2.41) على n نحصل على:

$$\bar{Y} = \hat{a} + \hat{b}\bar{X} \quad (2.42)$$

حيث \bar{X} و \bar{Y} هما متوسطا العينة، لكل من X و Y . وطالما أن \bar{X} و \bar{Y} معروفان من العينة، فإننا يكون لدينا معادلة ذات مجهولين، أي \hat{a} و \hat{b} وباللغة الفنية للاقتصاد القياسي فإن المعادلة (2.42) أو (2.41) تعرف «بالمعادلة الطبيعية». وقبل الاستمرار في تفسير هذه المعادلة واشتقاق معادلة أخرى، قد يكون من المفيد أن نشق هذه المعادلة الطبيعية بديها. ولنقوم بهذه المهمة، دعنا نرجع إلى نموذج الانحدار الأساسي.

$$Y_t = a + bX_t + u_t$$

فلو قمنا بتجميع الطرفين الأيمن والأيسر لهذه المعادلة عبر جميع القيم المشاهدة n للمتغيرين Y و X ، ثم، قسمنا كلا من الطرفين على n نحصل على:

$$\frac{\sum Y_t}{n} = \frac{\sum a}{n} + \frac{\sum bX_t}{n} + \frac{\sum u_t}{n} \quad (2.43)$$

التي يمكن تبسيطها إلى الصورة:

$$\bar{Y} = a + b\bar{X} + \frac{\sum u_t}{n} \quad (2.44)$$

ومن الافتراض (2a) في نموذج الانحدار، نعرف أن:

$$E(u_t) = 0$$

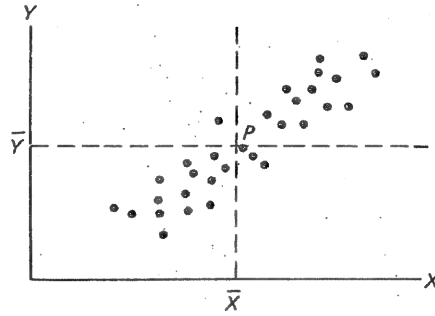
وهذا يعني أن القيمة المتوقعة للحد الأخير في (2.44) سوف تكون صفراً. ويتعين ملاحظة أن هذا لا يعني أن $\sum u_t/n$ سوف تساوي الصفر. وعموماً، فليس من المحتمل أن تساوي صفراً بالضبط، هذا على الرغم من أنه كلما كبر حجم العينة تناقص احتمال انحرافها عن الصفر بأي مقدار. وتنحو طريقة المتغير المساعد تجاه إهمال الحد $\sum u_t/n$ في (2.44) لأن قيمته المتوقعة تساوي صفراً. ولو تجاهلنا هذا الحد (أي افترضنا أنه يساوي صفراً) فإننا نحصل على:

$$\bar{Y} = \hat{a} + b\bar{X} \quad (2.45)$$

التي هي متماثلة مع (2.42). ويتعين ملاحظة أنه بالانتقال من (2.44) إلى (2.45) فقد تم إحلال \hat{a} و \hat{b} بدلا من a و b . والسبب في ذلك هو أن العلاقة المعبر عنها في المعادلة الطبيعية (2.45) تتسق مع (2.44) فقط في حالة أن يكون $\sum u_t/n = 0$. ومن ثم، فإنه إذا تحقق هذا الشرط فإن $\hat{a} = a$ و $\hat{b} = b$.

ولكن عموماً طالما أن $\sum u_t/n$ لا يساوي صفر بالضبط فإن \hat{a} و \hat{b} لا يساويان a و b . وأن كان \hat{a} و \hat{b} يعتبران مقدرين لـ a و b . ولعل الشئ التالي الذي توضحه المعادلة الطبيعية للعلاقة المقدرة بين Y و X ، هو أن الخط المثل لانتشار النقاط لا بد أن يمر بالنقطة التي يمثل محورها متوسطي المتغيرين بالعينة. وباستخدام الشكل (٢-١٣) فإن المعادلة الطبيعية توضح أن النقطة P تقع على الخط المقدّر. ويمكن أن

نرى الآن، بوضوح، الدور الذي يؤديه افتراضنا بأن $E(u_t) = 0$. فهذا الافتراض يسمح لنا بأن نرصد نقطة [أي $P(\bar{X}, \bar{Y})$] على الخط الذي سوف ن قدره من انتشار النقاط. وحيث إن أي نقطتين يحددان خطا مستقيما، فمن الواضح أنه إذا أمكننا إيجاد نقطة إضافية فسوف يصبح بمقدورنا تحديد معادلة للخط، كما سنحصل على علاقة مقدرة بين X و Y .



شكل (٢-١٣)

ولعمل ذلك، لابد من استخدام افتراض آخر، ويتمثل هذا في الافتراض الثالث بنموذج الانحدار والذي ينص على أن الخطأ العشوائي u_t يتعين أن يكون مستقلا عن X_t حيث: $cov(u_t, X_t) = 0$. وقد وضعنا أن هذا يتضمن:

$$E(u_t X_t) = 0$$

ويقودنا هذا إلى توقع بديهي وهو لو أن لدينا عينة من المشاهدات عن u_t و X_t فإن التغير المقدّر بينهما الذي توضحه المعادلة التالية:

$$\hat{\sigma}_{X,u} = \frac{\sum(u_t X_t)}{n}$$

سوف يساوي الصفر تقريبا. وحيث إن $E(u_t, X_t) = 0$ يتضمن أن $E(\hat{\sigma}_{X,u}) = 0$ فإن هذا يوضح أن الشرط الثاني الذي يمكن أن نفرضه على \hat{u}_t هو:

$$\frac{\Sigma(\hat{u}_t X_t)}{n} = 0 \quad (2.46)$$

أو: بضرب الطرفين في n :

$$\Sigma(\hat{u}_t X_t) = 0 \quad (2.47)$$

وبالعودة مرة أخرى للمعادلة (2.38):

$$Y_t = \hat{a} + \hat{b}X_t + \hat{u}_t$$

وبضرب طرفي المعادلة (2.38) في X_t نحصل على:

$$X_t Y_t = \hat{a}X_t + \hat{b}X_t^2 + \hat{u}_t X_t \quad (2.48)$$

وبجمع طرفي المعادلة (2.48) عبر كل القيم المشاهدة n للمتغيرين Y, X والقسمة على نحصل على:

$$\begin{aligned} \frac{\Sigma(X_t Y_t)}{n} &= \frac{\Sigma(\hat{a}X_t)}{n} + \frac{\Sigma(\hat{b}X_t^2)}{n} + \frac{\Sigma(\hat{u}_t X_t)}{n} \\ &= \hat{a}\bar{X} + \hat{b}\frac{\Sigma X_t^2}{n} + \frac{\Sigma(\hat{u}_t X_t)}{n} \end{aligned} \quad (2.49)$$

وبتطبيق الشرط $\Sigma(\hat{u}_t X_t) = 0$ على قيم العينة فإن هذا يتضمن أن الحد الأخير من (2.49) وهذا يعطينا:

$$\frac{\Sigma(X_t Y_t)}{n} = \hat{a}\bar{X} + \hat{b}\frac{\Sigma X_t^2}{n} \quad (2.50)$$

ويتوافر لدينا بذلك علاقة ثانية بين القيم المشاهدة للمتغيرين Y, X وقيم \hat{a} و \hat{b} التي سوف تحدد فيما بعد. وتمثل هذه العلاقة المعاداة الطبيعية الثانية.

ونظرا للأهمية المركزية للمعادلتين الطبيعييتين، فقد يكون من المفيد أن نستخدم مرة أخرى مدخلا أكثر بديهية لهذه العلاقة. وبالبدا مرة أخرى بعلاقة الانحدار الأساسية:

$$Y_t = a + bX_t + u_t$$

وبضرب جانبي المعادلة في X_t ، نحصل على:

$$X_t Y_t = a X_t + b X_t^2 + u_t X_t$$

وبالجمع بالنسبة لكل المشاهدات والقسمة على n نحصل على:

$$\begin{aligned} \frac{\Sigma(X_t Y_t)}{n} &= \frac{\Sigma(a X_t)}{n} + \frac{\Sigma(b X_t^2)}{n} + \frac{\Sigma(u_t X_t)}{n} \\ &= a \bar{X} + \frac{b \Sigma X_t^2}{n} + \frac{\Sigma(\hat{u}_t X_t)}{n} \end{aligned} \quad (2.51)$$

ونحن نعرف من افتراض سابق أن القيمة المتوقعة للحد الأخير في (2.51) تساوي الصفر. ولذا فإننا سوف نهمله بافتراض أن قيمته تساوي صفرًا. ومن ثم، تصبح المعادلة الطبيعية كمايلي:

$$\frac{\Sigma(X_t Y_t)}{n} = \hat{a} \bar{X} + \hat{b} \frac{\Sigma X_t^2}{n}$$

وجدير بالملاحظة أنه، بالانتقال من (2.51) التي تحتوي على المعلمتين a و b إلى المعادلة الطبيعية (2.50)، فإننا نكون قد أحللنا \hat{a} و \hat{b} محل a و b طالما أن $\Sigma(u_t X_t)/n$ لا يساوي بالضبط صفر. ويصبح أن تكون $\hat{a} = 0$ و $\hat{b} = 0$ فقط إذا كان $\Sigma(u_t X_t)/n$. ويتوافر لدينا الآن معادلتان (2.42)، (2.50) ومجهولان \hat{a} و \hat{b} . ومن ثم، تصبح في وضع يمكننا من القيام بالحل للحصول على المقدرين \hat{a} و \hat{b} ، ولعمل ذلك فمن المتعين أولاً أن نضرب (2.42) في X :

$$\bar{X} \bar{Y} = \hat{a} \bar{X} + \hat{b} \bar{X}^2 \quad (2.52)$$

ثم نطرح هذه المعادلة من (2.50) لنحصل على:

$$\frac{\Sigma(X_t Y_t)}{n} - \bar{X} \bar{Y} = \hat{b} \left(\frac{\Sigma X_t^2}{n} - \bar{X}^2 \right) \quad (2.53)$$

وبذلك نكون قد تخلصنا من \hat{a} وأصبح لدينا معادلة واحدة في مجهول واحد هو b . ويحل (2.53) بالنسبة لـ \hat{b} نحصل على:

$$\begin{aligned}\hat{b} &= \frac{[\Sigma(X_t Y_t) / n] - \bar{X}\bar{Y}}{[(\Sigma X_t^2 / n) - \bar{X}^2]} = \frac{\Sigma(X_t Y_t) - n\bar{X}\bar{Y}}{\Sigma X_t^2 - n\bar{X}^2} \\ &= \frac{\Sigma(X_t - \bar{X})(Y_t - \bar{Y})}{\Sigma(X_t - \bar{X})^2}\end{aligned}\quad (2.54)$$

وبعد حصولنا على \hat{b} ، يمكننا ببساطة استخدام (2.42) للحصول على \hat{a} :

$$\hat{a} = \bar{Y} - \hat{b}\bar{X} \quad (2.55)$$

وطالما أن هذا يعد أهم قسم في الكتاب، كما أنه أساسي لكل ما يأتي فيما بعد، فمن المفيد، في هذه المرحلة، أن نلخص ماتعرضنا له من قبل وأن نقدم مثالا رقميا بسيطاً. فلقد بدأنا بعلاقة خطية مفترضة بين متغيرين، ولم تكن هذه العلاقة مؤكدة وإنما كان مسموحاً فيها للمتغير التابع أن يتغير حول المتوسط عند كل قيمة من قيم المتغير المستقل. ولقد وصفنا طبيعة هذه العلاقة بنوع من التفصيل من خلال مجموعة من الافتراضات المتعلقة بخصائص التغيرات في قيمة المتغير التابع. وهذا يمثل ماهو معروف بنموذج الانحدار الخطي لمتغيرين أو الثنائي المتغيرات.

وتمثلت مشكلتنا في إيجاد طريقة لتقدير قيم معلمات هذه العلاقة. ولعمل ذلك تبيننا طريقة المتغير المساعد، التي وضعنا من خلالها مباشرة عددا من الشروط على مقدار الخطأ العشوائي. وعبرنا عن تلك الشروط من خلال افتراضات نموذج الانحدار نفسه. وعلى وجه التحديد لكي نحصل على \hat{a} و \hat{b} وضعنا شرطين هما $\Sigma \hat{u} / n = 0$ و $\Sigma(\hat{u}_t, X_t) / n = 0$. ونجم عن كل واحد من هذين الشرطين معادلة طبيعية. أو بمعنى آخر يمكننا كل شرط من تحديد موقع نقطة على الخط الذي يمثل انتشار النقاط المشاهدة، ومن خلال النقطتين، أمكننا أن نحل العلاقة المقدرة.

مثال

دعنا نستخدم الآن هذه الطريقة لتقدير علاقة اقتصادية باستخدام بيانات

الجدول (٢-٤) والذي يظهر مستويات الاستهلاك السنوي والدخل المتاح في الولايات المتحدة للسنوات ١٩٦٠-١٩٦٩ م. وربما تذكر أننا، في فحصنا للعلاقة بين الاستهلاك والدخل المتاح من قبل، ركزنا انتباهنا على مستويات الإنفاق الاستهلاكي للأسر ذات مستويات الدخل المختلفة. ومن خلال ما هو معروف بالتحليل القطاعي، استخدمنا معلومات عن ميزانيات عينة من الأسر عند نقطة زمنية معينة، ثم، بدأنا فحص كيفية تغير الإنفاق الاستهلاكي للأسر ذات الدخل المختلفة عند هذه النقطة من الزمن.

جدول (٢-٤) الإستهلاك والدخل المتاح في الولايات المتحدة (بليون دولار بالأسعار الجارية)

السنة	(C) الاستهلاك	(Y _d) الدخل التاح
١٩٦٠	٣٢٥	٣٥٠
١٩٦١	٣٣٥	٣٦٤
١٩٦٢	٣٥٥	٣٨٥
١٩٦٣	٣٧٥	٤٠٥
١٩٦٤	٤٠١	٤٣٨
١٩٦٥	٤٣٣	٤٧٣
١٩٦٦	٤٦٦	٥١٢
١٩٦٧	٤٩٢	٥٤٧
١٩٦٨	٥٣٧	٥٩٠
١٩٦٩	٥٧٦	٦٣٠

المصدر: التقرير الاقتصادي للرئيس، واشنطن، مكتب حكومة الولايات المتحدة للطباعة، فبراير ١٩٧٠ م،

ص ص ١٨٩-١٩٥.

وربما نكون قد درسنا مثلاً بيانات الدخل والاستهلاك للأسر عن عام ١٩٧٠ م، ومن ثم، فإن التحليل القطاعي يثبت الزمن. ومن المداخل البديلة استخدام تحليل السلاسل الزمنية، والذي نفحص، من خلاله، سلوك وحدة اقتصادية أو الوحدات ككل عبر الزمن. فعلى سبيل المثال، قد نفحص كيف يستجيب الإنفاق الاستهلاكي الكلي للمجتمع للتغير في الدخل

المتاح عبر الزمن. وهو ماسنفعله هنا مستخدمين بيانات تجميعية عن الولايات المتحدة الأمريكية. وعلى وجه التحديد، سوف نستخدم عشر مشاهدات عن الاستهلاك الكلي والدخل المتاح للولايات المتحدة في الجدول رقم (٢-٤) لتقدير أثر مستوى الدخل المتاح على الإنفاق الاستهلاك. ونفترض أولاً:

$$C_t = a + bY_{dt} + u_t$$

ثم نقدر قيمتي a و b ، حيث b يمكن تفسيرها على أنها الميل الحدي للاستهلاك. ولذا لا بد أن نحسب:

$$\hat{b} = \frac{\Sigma(C_t - \bar{C})(Y_{dt} - \bar{Y}_d)}{\Sigma(Y_{dt} - \bar{Y}_d)^2}$$

وأيضاً

$$\hat{a} = \bar{C} - \hat{b}\bar{Y}_d$$

وتظهر الحسابات اللازمة في الجدول (٢-٥)، وتصبح المعادلة المقدرة هي:

$$C = 13 + 0.89Y_d \quad (2.59)$$

ويظهر خط الانحدار AB المقدّر وانتشار النقاط العشرة في الشكل (٢-١٤)* ومن الواضح أن الخط يمدنا بتقريب جيد لنمط الاقتران بين C و Y_{dt} ، وهذا أمر سوف نتكلم عنه أكثر فيما بعد. ومن المثير للاهتمام، أيضاً، أن تتفق العلاقة المقدرة مع توقعاتنا النظرية المسبقة. فالقيمة المقدرة للميل الحدي للاستهلاك 0.89 وهي موجبة وتقع بين الصفر والواحد الصحيح، والقيمة المقدرة للحد الثابت هي 13 وهي موجبة أيضاً.

ولقد لاحظت من جدول (٢-٥) أن تحديد \hat{a} و \hat{b} قد تطلب قدراً كبيراً من الحسابات. وباستخدام بعض خصائص التجميع فإنه يمكن تخفيض هذه الحسابات لحد ما، وعلى وجه التحديد يلاحظ أن**

* عند هذه النقطة، تجاهل الخط CD في شكل (٢-١٤).

** انظر الافتراض الرابع في ملحق أ (A) من الفصل الأول.

جدول رقم (٢-٥)

(1)	(2)	(3)	(4)	(5)=(3)x(4)	(6)
C_t	Y_{dt}	$(C_t - \bar{C})$	$(Y_{dt} - \bar{Y}_d)$	$[(C_t - \bar{C})(Y_{dt} - \bar{Y}_d)]$	$(Y_{dt} - \bar{Y}_d)^2$
325	350	-105	-119	12,495	14,161
335	364	-95	-105	9,975	11,025
355	385	-75	-84	6,300	7,056
375	405	-55	-64	3,520	4,096
401	438	-29	-31	899	961
433	473	3	4	12	16
466	512	36	43	1,548	1,849
492	547	62	78	4,836	6,084
537	590	107	121	12,947	14,641
576	630	146	161	23,506	25,921

$$SC_t = 4,295 \quad \bar{C} = 430$$

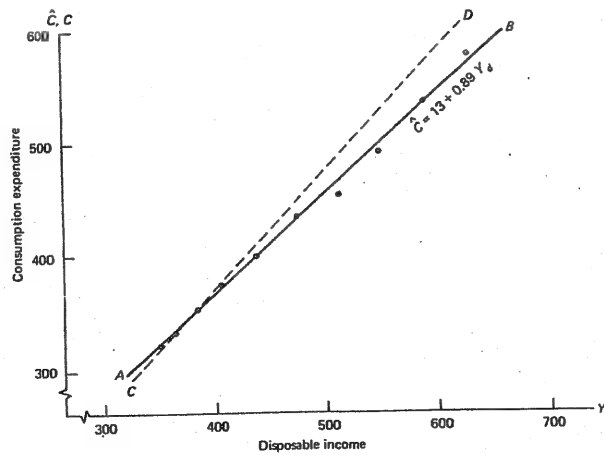
$$SY_{dt} = 4,694 \quad \bar{Y}_d = 469$$

$$\Sigma(C_t - \bar{C})(Y_{dt} - \bar{Y}_d) = 76,038$$

$$\Sigma(Y_{dt} - \bar{Y}_d)^2 = 85,810$$

$$\hat{b} = \frac{76,038}{85,810} = 0.89$$

$$\hat{a} = \bar{C} - \hat{b}\bar{Y}_d = 430 - 0.89(469) = 13$$



شكل (٢-١٤)

وبالمثل :

$$\Sigma(X_t - \bar{X})^2 = \Sigma(X_t - \bar{X})(X_t - \bar{X}) = \Sigma(X_t - \bar{X})X_t$$

وباستخدام هذه العلاقات، يمكننا تبسيط صيغة b لتبسيط الحسابات :

$$\begin{aligned} \hat{b} &= \frac{\Sigma(Y_t - \bar{Y})(X_t - \bar{X})}{\Sigma(X_t - \bar{X})^2} = \frac{\Sigma(Y_t - \bar{Y})X_t}{\Sigma(X_t - \bar{X})X_t} \\ &= \frac{\Sigma(Y_t X_t) - \bar{Y}\Sigma X_t}{\Sigma X_t^2 - \bar{X}\Sigma X_t} = \frac{\Sigma(Y_t X_t) - n\bar{Y}\bar{X}}{\Sigma X_t^2 - n\bar{X}^2} \end{aligned} \quad (2.57)$$

ولكن، حتى في هذه الصيغة، يتطلب \hat{a} و \hat{b} عملاً كثيراً، خاصة إذا كان هناك عدد كبير من المشاهدات، ولحسن الحظ، فإن هذه الحسابات يمكن إجراؤها بسهولة على الحاسوب، وهناك عدد كبير من البرامج المتاحة التي تنجز هذه الأعمال وتقدر قيم \hat{a} و \hat{b} بسهولة.

ملاحظة على أحد الافتراضات

قبل أن نبدأ في توضيح خصائص \hat{a} و \hat{b} يجب أن نتوقف برهة لتوضيح أهمية أحد الافتراضات الأساسية: وهو أنه يتعين أن تأخذ X الأقل، قيمتين مختلفتين. ولإثبات أهمية هذه النقطة، دعنا نتصور أن هذا الافتراض قد اختل بحيث إن X_t تأخذ قيمة معينة ولتكن X_0 وعندئذ، فإن المعادلتين الطبيعيين المستخدمتين في تحديد \hat{a} ، \hat{b} وهما (2.42) و (2.50) سوف يصبحان:

$$\bar{Y} = \hat{a} + \hat{b}X_0 \quad (2.42A)$$

and

$$X_0\bar{Y} = \hat{a}X_0 + \hat{b}X_0^2 \quad (2.50A)$$

وحيث إن $\Sigma X_t / n = X_0\bar{Y}$ و $\Sigma(X_t Y_t) / n = X_0\bar{Y}$ ، بقسمة (2.50A) على X_0 نحصل على:

$$\bar{Y} = \hat{a} + \hat{b}X_0$$

وتتماثل هذه المعادلة مع (2.42A). وكل هذا يعني أن لدينا معادلة واحدة هي (2.42A) ومجهولين هما \hat{a} و \hat{b} . ومن ثم، لانستطيع أن نحل النموذج للحصول على قيم وحيدة \hat{a} و \hat{b} ، كما لانستطيع أن نقدر a و b . والسبب البديهي لهذا هو أنه لو كانت مساوية دائماً X_0 فإن نموذج الانحدار:

$$Y_t = a + bX_t + u_t$$

سيصبح:

$$Y_t = a + bX_0 + u_t \quad (2.58)$$

وطالما أن X_0 ثابتة فإنها يمكن أن تدمج مع الحد الثابت a حيث يصبح النموذج:

$$Y_t = A + u_t \quad (2.59)$$

حيث $A = (a + bX_0)$ ، وبالتالي، فإن نموذج الانحدار يتراجع إلى نموذج يحتوي على حد ثابت وخطأ عشوائي.

ولو أردنا تقدير A فسوف نلجأ إلى طريقة المتغير المساعد مرة أخرى. وعلى وجه التحديد، سوف نلاحظ أولاً من المعادلة (2.59) أن:

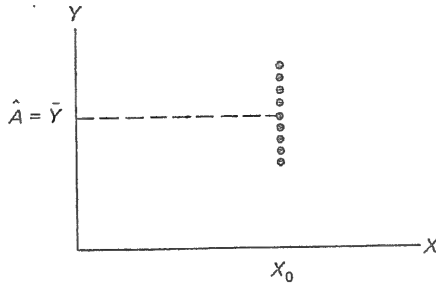
$$Y_t = \hat{A} + \hat{u}_t \quad (2.60)$$

ومن الافتراض $E(u_t) = 0$ ، نجد أن $\Sigma \hat{u}_t / n = 0$ ، وبالتالي، فإن المعادلة الطبيعية سوف تصبح:

$$\frac{\Sigma Y_t}{n} = \hat{A} \quad (2.61)$$

وبمعنى آخر، فإن مقدر A سوف يكون: $\hat{A} = \bar{Y}$ * ومن الشكل رقم (٢-١٥) يتضح أنه، إذا كانت X_t مساوية دائماً فإن شكل الانتشار سوف ينهار إلى سلسلة من النقاط المتراسة عمودياً فوق X_0 . ومن الواضح، أيضاً، أن هذه المجموعة من النقاط سوف تمكننا فقط من تقدير القيمة المتوسطة لـ Y و A المقابلة للقيمة المحددة لمتغير المستقل X و X_0 .

* لاحظ أنه، طالما لدينا معلمة واحدة فقط، A ، نريد تقديرها، فإننا نحتاج لمعادلة طبيعية واحدة، فقط.



شكل (٢-١٥)

ومن ثم، نجد أنه، لو لم تتغير قيمة X_t ، فإن طريقة المتغير المساعد لن تمكننا من تقدير أثر المتغير X على Y ، أي b ، منفصلاً عن قيمة الحد الثابت. وفي هذه الحال فإن كل ما يمكن عمله هو أن نحصل على مقدرة للأثر المزدوج $A = (a + bX_0)$. ومرة أخرى، فإن من البديهي لو أن X_t تأخذ دائماً قيمة واحدة، فإن أثرها على Y_t يصبح مختلطاً بالحد الثابت أو لا يمكن متصلة عنه. وسوف نستكمل هذه المناقشة في الفصل الرابع عندما نعمم هذه المشكلة في حالة الانحدار المتعدد.

(٢-٥) خواص \hat{a} و \hat{b}

يتوافر لدينا الآن طريقة للحصول على مقدرتي \hat{a} و \hat{b} . ولكن يتبقى لدينا التساؤل عما إذا كانت هذه الطريقة جيدة. بالطبع توجد هناك طرق أخرى للحصول على مقدرين لهاتين المعلمتين. فعلى سبيل المثال، يمكننا أن نأخذ أي نقطتين من انتشار النقاط بالشكل (٢-١٤)، ومن خلالهما، نشق خطاً يمكن استخدامه علاقة مقدرة بين الاستهلاك والدخل المتاح. ومن الواضح أن هذا سوف يكون أسهل بكثير من الإجراء الذي اتبعناه للوصول إلى المعادلتين (2.54) و (2.55). ولكن، بديهيًا ربما تشعر بأن الطريقة التي تبنيها هي أفضل الاثنين ذلك لأنها تستخدم كمية أكبر من المعلومات بانتظام. فالحظ AB الذي وفقناه من شكل الانتشار رقم (٢-١٤) يبدو أنه يصف بصورة معقولة السلوك الذي تعرضه المشاهدات. وعلى العكس من ذلك، إذا استخدمنا نقطتين فقط، يمكننا الحصول

على خط مثل CD في الشكل (٢-١٤) والذي يبدو أنه مقدر أقل جودة للعلاقة النمطية بين C و Y_d .

وسوف نوضح الآن أن \hat{a} و \hat{b} مقدرين جيدين، بمعنى أنهما يتمتعان بخصائص إحصائية معينة مرغوب فيها. وعلى وجه التحديد سوف نوضح أن:

١- القيمة المتوقعة لكل من \hat{a} و \hat{b} هي a و b على التوالي.

٢- تباين \hat{a} و \hat{b} صغير نسبياً.

ونتيجة لذلك، فسوف نعرف، في الأقل، أن مقدرنا موجهان إلى الهدف الصحيح وأن هامش الخطأ لهما صغير نسبياً بالمقارنة بالمقدرات الأخرى.

عدم التحيز*

سوف نثبت أولاً أن القيم المتوقعة لكل من \hat{a} و \hat{b} هي في الحقيقة a, b على التوالي، أو بمعنى آخر أن \hat{a} و \hat{b} مقدران غير متحيزين. وسوف نستخدم في الاثبات خمس خصائص لعملية الجمع.**

$$\Sigma(X_t - \bar{X}) = 0 \quad (2.62)$$

$$\Sigma(X_t + Y_t) = \Sigma X_t + \Sigma Y_t \quad (2.63)$$

$$\Sigma(X_t - \bar{X})(Y_t - \bar{Y}) = \Sigma(X_t - \bar{X})Y_t \quad (2.64)$$

$$\Sigma(X_t - \bar{X})(Y_t - \bar{Y}) = \Sigma(Y_t - \bar{Y})X_t \quad (2.65)$$

$$\Sigma(X_t - \bar{X})^2 = \Sigma(X_t - \bar{X})(X_t - \bar{X}) = \Sigma(X_t - \bar{X})X_t \quad (2.66)$$

* تتبع معالجة عدم التحيز تلك المعالجة الخاصة بجونستون:

J. Johnston, *Econometric Methods*, 2nd ed. New York: McGraw-Hill, 1972, pp. 18-20.

** اثبتنا كل هذه الخصائص رياضياً في الملحق أ (A) في الفصل الأول.

ومن صيغة \hat{b} في المعادلة (2.54)، نجد أن:

$$\hat{b} = \frac{\Sigma(X_t - \bar{X})(Y_t - \bar{Y})}{\Sigma(X_t - \bar{X})^2}$$

وباستخدام (2.64)، يمكن تبسيط هذه الصيغة إلى:

$$\hat{b} = \frac{\Sigma(X_t - \bar{X})Y_t}{\Sigma(X_t - \bar{X})^2} \quad (2.67)$$

ولو أننا أحللنا الآن $Y_t = a + bX_t + u_t$ في بسط المعادلة (2.67)، نحصل على:

$$\hat{b} = \frac{\Sigma(X_t - \bar{X})(a + bX_t + u_t)}{\Sigma(X_t - \bar{X})^2} \quad (2.68)$$

وبتوسيع بسط المعادلة (2.68) واستخدام الخاصية الموجودة في المعادلة (2.63)، نحصل على:

$$\hat{b} = \frac{a\Sigma(X_t - \bar{X}) + b\Sigma(X_t - \bar{X})X_t + \Sigma(X_t - \bar{X})u_t}{\Sigma(X_t - \bar{X})^2} \quad (2.69)$$

دعنا الآن نكتب المعادلة (2.69) على النحو:

$$\hat{b} = \frac{a\Sigma(X_t - \bar{X})}{\Sigma(X_t - \bar{X})^2} + \frac{b\Sigma(X_t - \bar{X})X_t}{\Sigma(X_t - \bar{X})^2} + \frac{\Sigma(X_t - \bar{X})u_t}{\Sigma(X_t - \bar{X})^2} \quad (2.70)$$

ومن (2.62)، يتضح أن الحد الأول في الطرف الأيمن بالمعادلة (2.70) يساوي

الصفر، وباستخدام (2.66) لتغيير صيغة المقام في الحد الثاني بالطرف الأيمن في المعادلة

(2.70) إلى $\Sigma(X_t - \bar{X})X_t$ ، نجد أن الثاني، ببساطة، يساوي b . ومن ثم، فإن:

$$\hat{b} = b + \frac{\Sigma(X_t - \bar{X})u_t}{\Sigma(X_t - \bar{X})^2} \quad (2.71)$$

ولتبسيط التحليل التالي، دعنا نستخدم الرموز التالية:

$$A = \sum (X_t - \bar{X})^2 \quad (2.72)$$

و

$$w_t = (X_t - \bar{X}) \quad (2.73)$$

وباستخدام هذه التعريفات، فإن صيغة \hat{b} في (2.71) تصبح:

$$\begin{aligned} \hat{b} &= b + \frac{\sum w_t u_t}{A} = b + \frac{w_1 u_1}{A} + \frac{w_2 u_2}{A} + \dots + \frac{w_n u_n}{A} \\ &= b + \left(\frac{w_1}{A}\right) u_1 + \left(\frac{w_2}{A}\right) u_2 + \dots + \left(\frac{w_n}{A}\right) u_n \end{aligned} \quad (2.74)$$

ونصبح في وضع الآن يمكننا من إثبات أن \hat{b} غير متحيزة. وعلى وجه التحديد، من (2.74) لدينا:

$$E(\hat{b}) = b + E\left[\left(\frac{w_1}{A}\right) u_1\right] + E\left[\left(\frac{w_2}{A}\right) u_2\right] + \dots + E\left[\left(\frac{w_n}{A}\right) u_n\right] \quad (2.75)$$

وطالما أن الحدود $(w_1/A), \dots, (w_n/A)$ تعتمد، فقط، على الـ n قيمة للمتغير المستقل X_t ، وأن قيم المتغير المستقل وقيم الخطأ العشوائي يفترض أنها مستقلة عن بعضها، يصبح لدينا:

$$E(\hat{b}) = b + E\left(\frac{w_1}{A}\right) E(u_1) + E\left(\frac{w_2}{A}\right) E(u_2) + \dots + E\left(\frac{w_n}{A}\right) E(u_n) \quad (2.76)$$

ونحن نعرف من نموذج الانحدار أن $E(u_t) = 0$. ولهذا، فإن القيم المتوقعة لكل الحدود ماعدا الأول تساوي الصفر، ومن ثم:

$$E(\hat{b}) = b \quad (2.77)$$

ولذا، فإن \hat{b} هي مقدر غير متحيز لـ b .

وبالتحول الآن إلى \hat{a} ، فإن لدينا من (2.55):

$$\hat{a} = \bar{Y} - \hat{b}\bar{X} \quad (2.55)$$

وطالما أن $Y_i = a + bX_i + u_i$ فإنه يتبع ذلك :

$$\bar{Y} = a + b\bar{X} + \bar{u} \quad (2.78)$$

وبإحلال (2.78) في (2.55)، نحصل على :

$$\hat{a} = a + b\bar{X} + \bar{u} - \hat{b}\bar{X} \quad (2.79)$$

وبإحلال (2.74) بدلا من \hat{b} في (2.79)، نحصل على :

$$\hat{a} = a + b\bar{X} + \bar{u} - b\bar{X} - \left(\frac{w_1 \bar{X}}{A} \right) u_1 - \left(\frac{w_2 \bar{X}}{A} \right) u_2 - \dots - \left(\frac{w_n \bar{X}}{A} \right) u_n \quad (2.80)$$

وبملاحظة أن $b\bar{X}$ تسقط مع $b\bar{X}$ - وبأخذ القيمة المتوقعة لـ \hat{a} نحصل على :

$$E(\hat{a}) = a + E(\bar{u}) - E\left(\frac{w_1 \bar{X}}{A} \right) E(u_1) - E\left(\frac{w_2 \bar{X}}{A} \right) E(u_2) - E\left(\frac{w_n \bar{X}}{A} \right) E(u_n) = a \quad (2.81)$$

وحيث إن $E(u_i) = 0$ ، $E(\bar{u}) = 0$ فإن \hat{a} أيضا، يكون مقدر غير متحيز .

تباينات \hat{a} و \hat{b} : بعض الأساسيات

تبقى قضية تباين كل من \hat{a} و \hat{b} . ونحن نعرف الآن أن طريقة المتغير المساعد قد أنتجت مقدرين قيمتا وسطيهما هما قيمتا المعلمتين المقابلتين لهما . والسؤال الذي يبرز الآن هو عن المدى المتوقع أن تنحرف به \hat{a} و \hat{b} عن قيمتي الوسطين a و b . ونأمل بالطبع أن يتمخض اجراءونا عن مقدرين تباينهما أقل نسبيا من تباين المقدرات التي تتولد عن طرق أخرى . وقبل أن نشق صيغتي التباين لكل من \hat{a} و \hat{b} ، فقد يكون من المفيد أن نناقش ، باختصار ، لماذا يوجد هناك تباين لهذين المقدرين في المقام الأول .

دعنا نبدأ بفرض وجود عيتين تتكون كل واحدة منهما من خمس مشاهدات

عن X و Y :

عينة (٢)	عينة (١)	
$(Y_{12} \quad Y_1)$	$(Y_{11} \quad X_1)$	مشاهدة ١
$(Y_{22} \quad X_2)$	$(Y_{21} \quad X_2)$	مشاهدة ٢
$(Y_{32} \quad X_3)$	$(Y_{31} \quad X_3)$	مشاهدة ٣
$(Y_{42} \quad X_4)$	$(Y_{41} \quad X_4)$	مشاهدة ٤
$(Y_{52} \quad X_5)$	$(Y_{51} \quad X_5)$	مشاهدة ٥

ويشير الرقم الأول في الدليل السفلي للمتغير Y إلى المشاهدة، في حين يشير الرقم الثاني إلى العينة. وفي مثالنا هذا، نفترض أن مجموعة قيم المتغير X واحدة في العيتين. وبالطبع، لا يتضمن هذا الافتراض الخاص بقيم المتغير X أن قيم المتغير Y سوف تكون متساوية في العيتين. والسبب في ذلك هو أن قيم المتغير Y تعكس أثرين: (١) أثر قيمة X التي تعمل من خلال العلاقة المتوسطة $Y^m = a + bX$ ، و(٢) أثر حد الخطأ u الذي هو متغير عشوائي متوسطه صفر ويفترض فيه أنه مستقل عن المتغير X .

دعنا الآن نفحص المشاهدة الأولى في كل عينة من العيتين. ففي غياب الخطأ العشوائي، تتساوي Y_{11} مع Y_{12} طالما أن Y في كل حالة تعكس أثر X_1 فقط. وفي هذا المثال تأخذ كل من Y_{11} و Y_{12} القيمة $Y = a + bX_1$. ولو كان هذا صحيحا لكل المشاهدات فإنه من الواضح أن مجموعتي القيم المشاهدة لكل من X و Y سوف تكون متطابقة في العيتين، ومن ثم، فإن القيمتين اللتين يمكن حسابهما لكل من \hat{a} و \hat{b} سوف تتساويان في كل حالة مع القيمتين الفعليتين لـ a و b . غير أن ظهور خطأ عشوائي يتضمن أن القيمة المشاهدة للمتغير Y سوف تنحرف لحد ما عن القيمة التي تعكس أثر X فقط. وعلى وجه التحديد:

$$Y_{11} = a + bX_1 + u_{11}$$

$$Y_{12} = a + bX_1 + u_{12}$$

حيث $u_{11} \neq u_{12}$ عموماً. وهذا يعد صحيحاً، أيضاً، للقيم المشاهدة الأخرى للمتغير Y ، حيث نجد، عموماً، أن $Y_{11} \neq Y_{12}$. وطالما أن \hat{a} و \hat{b} تحسبان مباشرة من القيم المشاهدة X و Y ، فمن الواضح، عموماً أن \hat{a}_1 و \hat{b}_1 لن تساوي \hat{a}_2 و \hat{b}_2 على التوالي، حيث يشير الدليل السفلي إلى رقم العينة التي قدر \hat{a} و \hat{b} منها. ومن ثم، فإن الخطأ العشوائي سوف يؤدي إلى اختلاف القيمة المشاهدة لـ Y وبالتالي القيم المحسوبة لـ \hat{a} و \hat{b} من عينة لأخرى.

دعنا الآن نعمم ماسبق. افترض أن لدينا عدد P عينة مكونة من عدد محدد من المشاهدات عن X و Y ، وأن مجموعة القيم الخاصة بالمتغير المستقل X نفسها في كل العينات، وأفترض، أيضاً أن \hat{a}_i و \hat{b}_i هي قيمتي \hat{a} و \hat{b} المحسوبتان من العينة i . وطالما أن قيم Y تختلف من عينة لأخرى. ومن ثم، فإنه في ظل شروط عامة، لو أن P كانت لانهائية (أي إذا كان هناك عدد لانهاية من العينات) فإن المجاميع (A) و (B) في (2.82) سوف تساوي تباين \hat{a} و σ_a^2 ، وتباين \hat{b} و σ_b^2 ، وباحتمال يساوي الواحد الصحيح:

$$\frac{\sum_{i=1}^P (\hat{a}_i - a)^2}{P} \quad (2.82A)$$

$$\frac{\sum_{i=1}^P (\hat{b}_i - b)^2}{P} \quad (2.82B)$$

وسوف نشق صيغاً خاصة بتباين كل من \hat{a} و \hat{b} في القسم التالي، وحتى هذه اللحظة يتعين ملاحظة أن التباينات المشار إليها أعلاه تسمى التباينات المشروطة. أي أن صيغ التباينات السابقة بنيت على أساس افتراض أن مجموعة قيم X هي نفسها عبر كل العينات. ومن ثم، فإن اختلاف قيم \hat{a} و \hat{b} عبر العينات ترجع تماماً إلى اختلاف قيم الخطأ العشوائي من عينة لأخرى. وبالطبع، كما يتوقع المرء أن

قيم الصيغ الموضحة في المعادلة (2.82) تعتمد جزئياً على قيم X عموماً. وفي الواقع، فإن قيمة التباين المشروط تعطي الباحث مؤشراً عن درجة عدم التأكد الخاصة بالمقدر الذي بنى على بيانات معينة للمتغير X . وفي النهاية، فإننا نلاحظ أنه، لكون $E(\hat{a}) = a$ و $E(\hat{b}) = b$ فسوف تتحقق الصيغ التالية تحت شروط عامة:

$$\sum_{i=1}^P \frac{\hat{a}_i}{P} = a \quad \text{و} \quad \sum_{i=1}^P \frac{\hat{b}_i}{P} = b \quad (2.83)$$

باحتمال يساوي الواحد الصحيح.

تباين المقدرات*

سوف نشق الآن صيغتين أولهما لتباين \hat{b} وثانيهما لتباين \hat{a} . ولنبدأ أولاً بالصيغة الأساسية لـ \hat{b} :

$$\hat{b} = \frac{\Sigma(X_t - \bar{X})(Y_t - \bar{Y})}{\Sigma(X_t - \bar{X})^2}$$

ولقد أوضحنا في إثباتنا لخاصية عدم التحيز أن:

$$\hat{b} = b + \frac{w_1 u_1}{A} + \frac{w_2 u_2}{A} + \dots + \frac{w_n u_n}{A} \quad (2.74)$$

حيث:

$$w_t = (X_t - \bar{X}) \quad \text{و} \quad A = \Sigma(X_t - \bar{X})^2$$

وحتى نستخدم (2.74) في اشتقاق صيغة لتباين \hat{b} و $\text{var}(\hat{b})$ ، فنحن في حاجة لاستخدام علاقة أساسية عن تباين مجموع توليفة خطية من متغيرات عشوائية.

* سوف نبسط التعبير الذي نستخدمه، أحياناً، باستعمال تباينات \hat{a} و \hat{b} ، ويتعين على القارئ أن يتذكر أن

هذه تباينات مشروطة. أرجع إلى: J. Johnston, *Econometric Methods*, 2nd ed., 1972, pp. 18-20. لاشتقاق مماثل للنتائج الواردة نفسها في هذا القسم

وعلى وجه التحديد (هذه الصيغة مشتقة في ملحق الفصل الثاني) لو أن لدينا متغيراً عشوائياً M معرف على النحو التالي:

$$M = a_0 + a_1 Z_1 + a_2 Z_2 + \dots + a_n Z_n \quad (2.84)$$

حيث إن الـ a 's ثوابت، والـ Z 's هي متغيرات عشوائية، فبافتراض أن Z 's غير مرتبطة فإن:

$$\text{var}(M) = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \dots + a_n^2 \sigma_n^2 \quad (2.85)$$

حيث:

$$\sigma_j^2 = \text{var}(Z_j)$$

وتصبح الآن أهمية مفهوم التباين المشروط واضحة. فعلى وجه التخصيص، لو أننا مهتمون بتباين \hat{b} المقابل لمجموعة من قيم X المعطاة عبر عدد من العينات، فإن العدد n من قيم X يمكن اعتبارها، ببساطة، ثوابت. ونستخلص من (2.74) أن الـ n قيمة لـ w_t وقيمة A يمكن اعتبارها ثوابت كذلك. وفي ظل هذه الشروط، فإن \hat{b} في (2.74) تصبح ببساطة توليفة خطية من أخطاء عشوائية. وطالما أن هذه الأخطاء العشوائية مستقلة، ومن ثم، غير مرتبطة مع بعضها بعضاً افتراضاً، كما أن لها التباين σ_u^2 نفسه، يصبح لدينا، بتطبيق (2.82):

$$\begin{aligned} \text{var}(\hat{b}) &= \sigma_{\hat{b}}^2 = \frac{w_1^2 \sigma_u^2}{A^2} + \frac{w_2^2 \sigma_u^2}{A^2} + \dots + \frac{w_n^2 \sigma_u^2}{A^2} \\ &= \frac{\sigma_u^2}{A^2} \sum w_t^2 = \sigma_u^2 \frac{\sum (X_t - \bar{X})^2}{\left[\sum (X_t - \bar{X})^2 \right]^2} \end{aligned} \quad (2.86)$$

ومن ثم، تكون نتيجتنا:

$$\text{var}(\hat{b}) = \sigma_u^2 = \frac{\sigma_u^2}{\sum (X_t - \bar{X})^2} \quad (2.87)$$

وتمثل الصيغة (2.87) تباين \hat{b} المقابل لأي مجموعة معينة من قيم X . وليس من

المفاجئ أن نجد أن تباين المقدّر \hat{b} يتغير مباشرة مع تباين الخطأ العشوائي، فالأي مجموعة من قيم X_i ، كلما زاد تباين الخطأ العشوائي كلما زاد تباين \hat{b} . وبديها، كلما زاد عدم التأكد في نموذج الانحدار الأساسي، انخفضت الثقة في المقدّر الذي نحصل عليه.

وبالمثل، يمكن أن نشق تباين \hat{a} ، فالصيغة الأصلية لـ \hat{a} هي:

$$\hat{a} = \bar{Y} - \hat{b}\bar{X} \quad (2.55)$$

وبملاحظة أن $\bar{Y} = a + b\bar{X} + \bar{u}$ ، يمكن إحلال (2.74) بدلا من \hat{b} في (2.55) لنحصل على:

$$\begin{aligned} \hat{a} &= a + b\bar{X} + \bar{u} - b\bar{X} - \left(\frac{\bar{X}w_1}{A}\right)u_1 - \dots - \left(\frac{\bar{X}w_n}{A}\right)u_n \\ &= a + \bar{u} - \left(\frac{\bar{X}w_1}{A}\right)u_1 - \dots - \left(\frac{\bar{X}w_n}{A}\right)u_n \end{aligned} \quad (2.88)$$

وبالتعبير عن \bar{u} كمايلي:

$$\bar{u} = \frac{u_1}{n} + \dots + \frac{u_n}{n} \quad (2.89)$$

وإحلالها في (2.88) مع دمج الحدود نحصل على:

$$\hat{a} = a + \gamma_1 u_1 + \dots + \gamma_n u_n \quad (2.90)$$

حيث:

$$\gamma_1 = \left[(1/n) - \bar{X}w_1 / A \right]$$

ومن ثم، يتضح أنه، في ظل الافتراضات المتعلقة بقيم X ، فإن \hat{a} تتناقص إلى توليفة خطية من أخطاء عشوائية. ولهذا فإنه بتطبيق (2.85) نجد أن تباين \hat{a} يصبح:

$$\begin{aligned} \text{var}(\hat{a}) &= \sigma_a^2 = \gamma_1^2 \sigma_u^2 + \dots + \gamma_n^2 \sigma_u^2 \\ &= \sigma_u^2 \sum_{t=1}^n \gamma_t^2 \end{aligned} \quad (2.91)$$

والآن، لاحظ أن:

$$\begin{aligned}\Sigma Y_t^2 &= \Sigma \left[\frac{1}{n^2} + \left(\frac{\bar{X}^2}{A^2} \right) w_t^2 - \left(2 \frac{\bar{X}}{nA} \right) w_t \right] \\ &= \frac{1}{n} + \left(\frac{\bar{X}^2}{A^2} \right) \Sigma w_t^2 - \left(\frac{2\bar{X}}{nA} \right) \Sigma w_t\end{aligned}\quad (2.92)$$

وحيث إن:

$$\Sigma w_t^2 = \Sigma (X_t - \bar{X})^2 = A \quad (2.93)$$

و

$$\Sigma w_t = \Sigma (X_t - \bar{X}) = 0 \quad (2.94)$$

فإن تبين \hat{a} كما هو معطى في (2.91)، يمكن التعبير عنه كمايلي:

$$\begin{aligned}\sigma_a^2 &= \sigma_u^2 \left(\frac{1}{n} + \frac{\bar{X}^2}{A} \right) \\ &= \sigma_u^2 \left(\frac{A + n\bar{X}^2}{nA} \right)\end{aligned}\quad (2.95)$$

وأخيرا، نجد من الملحق A في الفصل الأول أن:

$$\Sigma (X_t - \bar{X})^2 = \Sigma X_t^2 - n\bar{X}^2$$

وبالإحلال مباشرة في (2.95)، نحصل على:

$$\sigma_a^2 = \frac{\sigma_u^2 \Sigma X_t^2}{n \Sigma (X_t - \bar{X})^2} \quad (2.96)$$

وكما هو في حالة σ_b^2 نجد أنه، لأي مجموعة معطاة من قيم X_t ، فإن قيمة σ_a^2

تتغير بصورة مباشرة مع تبين الخطأ العشوائي σ_u^2 .

وقد يكون من المفيد من هذا المنطلق أن نقوم بحساب التباينات الخاصة بـ \hat{a}

و \hat{b} لعينة افتراضية من قيم X_t ولـ σ_u^2 . افترض أننا بناءً على معلومات أخرى نعرف أن تباين الخطأ العشوائي σ_u^2 يساوي عشرة، وافترض أن لدينا مجموعة من القيم المشاهدة لكل من X و Y كما هو موضح بالجدول رقم (٦-٢)

جدول رقم (٦-٢)	
Y	X
8	3
12	6
14	10
15	12
15	14
18	15

ومن ثم، فإنه، في هذه الحالة:

$$\sum X_t^2 = 710, \quad \sum (X_t - \bar{X})^2 = 110, \quad n = 6$$

وباستخدام التعبيرات (2.87) و (2.96)، نجد أنه لهذه المجموعة من القيم الخاصة بـ X_t ، يوجد لدينا:

$$\text{var}(\hat{b}) = \frac{10}{110} = 0.09, \quad \text{var}(\hat{a}) = \frac{10(710)}{6(110)} = 10.8$$

خاصية أصغر تباين

يتوافر لدينا الآن صيغتان لتباين كل من \hat{a} و \hat{b} وسوف تتضح أهميتهما في تقدير فترات الثقة عندما نأتي لمشكلة اختبار الفرضيات. ولكن قبل ذلك نريد أن نعرف أولاً ما إذا كان هذان التباينان أكبر أم أصغر للتباينات المصاحبة لطرق تقدير أخرى لـ a و b . وفي هذا الصدد يمكن إثبات أنه لا توجد مقدرات خطية غير متحيزة لـ a و b تتميز بتباينات أقل من تباين كل من \hat{a} و \hat{b} التي اشتقت في هذا الفصل. ونقصد بالمقدرات الخطية تلك المقدرات التي يمكن التعبير عنها بوصفها توليفات خطية من قيم المتغير التابع Y . فعلى سبيل المثال، نتذكر من (2.67) أن:

$$\begin{aligned}\hat{b} &= \frac{\sum (X_t - \bar{X})Y_t}{\sum (X_t - \bar{X})^2} = \frac{\sum w_t Y_t}{A} \\ &= \left(\frac{w_1}{A}\right)Y_1 + \dots + \left(\frac{w_n}{A}\right)Y_n\end{aligned}\quad (2.97)$$

ومن (2.55)، نجد أن:

$$\begin{aligned}\hat{a} &= \bar{Y} - \bar{X}\hat{b} = \frac{\sum Y_t}{n} - \bar{X} \frac{\sum w_t Y_t}{A} \\ &= \sum \gamma_t Y_t = \gamma_1 Y_1 + \dots + \gamma_n Y_n\end{aligned}\quad (2.98)$$

وعلى الرغم من أن إثبات خاصية أقل تباينا ليس صعبا، إلا أنه طويل لحد ما، ولهذا السبب، فقد اخترنا أن نضعه في ملحق بنهاية هذا الفصل. ونحن نشجع القارئ على محاولة تتبع خطوات الإثبات بنفسه، ولكن، إذا فضلت أن تأخذ هذا القول اعتقادا (في الأقل في الظروف الحالية) فلن يسبب لك هذا أية مصاعب في فهم المادة العلمية القادمة.

والآن، لم يعد يتوافر لدينا طريقة لتقدير a و b فقط، بل لدينا من الأسباب ما يجعلنا نعتقد أنها طريقة جيدة. فأولا هي طريقة يتولد عنها مقدرات غير متحيزة للمعلمات، وثانيا تتمتع هذه المقدرات بخاصية أقل تباينا بين كل مجموعة المقدرات الخطية غير المتحيزة لكل من a و b .

مقدرات التباين

يوجد هناك معلومة إضافية لم ننته منها بعد تخص تباين \hat{a} و \hat{b} . فبالرغم من أن لدينا صيغتين خاصيتين بهما في المعادلتين (2.87) و (2.96) فإن هاتين الصيغتين تحتويان على σ_u^2 أي تباين الخطأ العشوائي في نموذج الانحدار. والمشكلة هنا أن σ_u^2 غير معروفة مثلها في ذلك مثل كل من a و b . وهذا يعني أنه، لكي نحصل على قيمتي التباين لكل من \hat{a} و \hat{b} فإنه يتعين علينا أولا أن نقدر σ_u^2 . وسوف نتابع الآن

مشكلة اشتقاق مقدر لـ σ_u^2 .

$$\sigma_u^2 = E(u_t - 0)^2 = E(u_t^2) \quad \text{تذكر أولاً أن:}$$

أي أن تباين الخطأ العشوائي يتمثل، ببساطة، في متوسط تربيعه. والآن، من نموذج الانحدار الأساس نعرف أن:

$$u_t = Y_t - a - bX_t = Y_t - Y_t^m \quad (2.35)$$

وافترض أننا عرفنا قيمتي a و b . وفي هذه الحالة لو أن لدينا عينة بحجم n من المشاهدات الخاصة بقيم X و Y يمكننا باستخدام (2.35)، اشتقاق عدد n من القيم لـ u_t . وعندئذ، سوف يكون من المعقول أن نقدر σ_u^2 عن طريق الحصول على القيمة المتوسطة لـ u_t^2 في العينة:

$$\frac{\sum u_t^2}{n} = \frac{\sum (Y_t - a - bX_t)^2}{n} \quad (2.99)$$

غير أن هذا لا يمكن عمله في الواقع، لأننا، عموماً، لا يمكننا معرفة قيمتي a و b . والجراء الواضح هو أن نحصل على مقدر لـ σ_u^2 وليكن $\hat{\sigma}_u^2$ ، من المعادلة (2.99) عن طريق إحلال \hat{a} و \hat{b} محل a و b . وعندئذ سوف يكون لدينا:

$$\hat{\sigma}_u^2 = \frac{\sum (Y_t - \hat{a} - \hat{b}X_t)^2}{n-2} = \frac{\sum (Y_t - \hat{Y}_t)^2}{n-2} = \frac{\sum \hat{u}_t^2}{n-2} \quad (2.100)$$

ويلاحظ أن مقام المعادلة (2.100) هو $(n-2)$ وليس n ، الأمر الذي يوضح (كما ناقشنا من قبل) أن لدينا، فقط، $(n-2)$ درجة حرية في البسط. ولقد فقدنا درجتين حرية لأننا أحللنا مقدرين محل معلمتين. وبالرغم من أننا لن نثبت ذلك هنا، إلا أن هذه الحالة تتضمن أن:

$$E(\hat{\sigma}_u^2) = \sigma_u^2 \quad (2.101)$$

حيث $\hat{\sigma}_u^2$ هو مقدر غير متحيز لـ σ_u^2 .

ويتوافر لدينا الآن مقدر لتباين الخطأ العشوائي، ومن ثم، فإن تباين \hat{a} و \hat{b} يمكن الحصول عليهما، ببساطة بإحلال $\hat{\sigma}_u^2$ محل σ_u^2 في المعادلتين المقابلتين (2.87) و (2.96) وعلى وجه التحديد فإن مقدري تباين كل من \hat{a} و \hat{b} يصبحان:

$$\hat{\sigma}_a^2 = \frac{\hat{\sigma}_u^2 \sum X^2}{n \sum (X_i - \bar{X})^2} \quad (2.102)$$

$$\hat{\sigma}_b^2 = \frac{\hat{\sigma}_u^2}{\sum (X_i - \bar{X})^2}$$

وأخيرا، فإنه نتيجة لأن $\hat{\sigma}_u^2$ غير متحيزة من المعادلة (2.101)، فإن $\hat{\sigma}_a^2$ و $\hat{\sigma}_b^2$ مقدران غير متحيزين.

مثال

لقد استخدمنا سابقا في جدول رقم (٢-٥) بيانات عن الاستهلاك والدخل المتاح للولايات المتحدة الأمريكية لحساب قيم \hat{a} و \hat{b} لدالة الاستهلاك المقدرة. وربما نتذكر أن المعادلة المقدرة كانت:

$$C = 13 + 0.89Y_d \quad (2.56)$$

ونحن الآن في وضع يمكننا فيه حساب القيم المقدرة للتباينات المقابلة. فأولا، بالإشارة إلى الجدول رقم (٢-٧) يمكننا حساب قيمة $\hat{\sigma}_u^2$:

$$\hat{\sigma}_u^2 = \frac{92}{10 - 2} = 11.5$$

ومن الجدول (٢-٥)، نجد أن:

$$\sum (Y_{dt} - \bar{Y}_d)^2 = 85,810$$

ومن ثم، نجد:

$$\Sigma(Y_{di}^2) = 2,289,172$$

وبالتالي:

$$\hat{\sigma}_a^2 = \frac{11.5(2,289,172)}{10(85,810)} = 31$$

$$\hat{\sigma}_b^2 = \frac{11.5}{85,810} = 0.0001$$

خاصية أصغر المربعات لكل من \hat{a} و \hat{b}

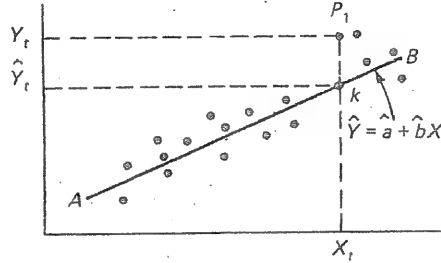
يوجد هناك خاصية أخيرة للمقشرين \hat{a} و \hat{b} نريد أن نشير إليها. فلكي نقدر العلاقة بين X و Y يوجد هناك مدخل آخر وهو أن نحاول توليف خط لنقاط الانتشار بحيث يكون اقرب ما يمكن إلى هذه النقاط. وافترض على سبيل المثال أنه علينا الآن اختيار \hat{a} و \hat{b} اللتين تجعلان الخط AB في الشكل (٢-١٦) يمثل نقاط الانتشار أفضل تمثيل.

جدول رقم (٢-٧)

Year	C	\hat{C}	$\hat{u} = C - \hat{C}$	$\hat{u}^2 = (C - \hat{C})^2$
1960	325	325	0	0
1961	335	337	-2	4
1962	355	356	-1	1
1963	375	373	2	4
1964	401	403	-2	4
1965	433	434	-1	1
1966	466	469	-3	9
1967	492	500	-8	64
1968	537	538	-1	1
1969	576	574	2	4

$$\Sigma(C_i - \hat{C})^2 = 92$$

بسبب تقريب الأرقام فإن $\Sigma \hat{u}_i$ ليس صفراً.



شكل رقم (٢-١٦)

وفي هذا الصدد، دعنا نأخذ نقطة مثل P_1 بشكل الانتشار (٢-١٦)، وبالطبع، فإنها لن تقع، عموماً، على الخط حرفياً، وإنما بسبب وجود الخطأ العشوائي u_t ، فإنها تقع أعلى أو أسفل الخط AB . ويلاحظ هنا أن المسافة الرأسية التي تنحرف بها النقطة عن الخط وهي P_1k في هذه الحالة، تمثل الفرق بين القيمة المشاهدة للمتغير Y (Y_t) والتي تقابل القيمة X_t ، والقيمة المحسوبة للمتغير نفسه Y وهي \hat{Y}_t ، حيث يمكن الحصول على \hat{Y}_t من العلاقة المقدرة الممثلة بالخط AB :

$$P_1k = (Y_t - \hat{Y}_t) = (Y_t - \hat{a} - \hat{b}X_t) \quad (2.103)$$

افترض الآن أننا نريد وضع الخط AB حيث يدني المسافة التي تمثل انحرافات القيم المشاهدة عنه. وإحدى المشاكل التي تواجهنا عند تدنية مجموع هذه الانحرافات هي أن $(Y_t - \hat{Y}_t) > 0$ للنقاط التي تقع فوق الخط، هذا في حين أن $(Y_t - \hat{Y}_t) < 0$ للنقاط التي تقع تحت الخط AB . ومن ثم، فإنه من الممكن أن يوجد هناك انتشار واسع جداً للنقاط حول الخط بالرغم من أن المجموع الجبري للانحرافات يكون صغير جداً وربما صفراً. وفي الحقيقة، يمكننا تدنية هذا المجموع حرفياً بوضع الخط AB في أعلى مستوى ممكن طالما أن هذا سوف ينتج عنه قيمة سالبة كبيرة للمجموع $\sum (Y_t - \hat{Y}_t)$. وإحدى الطرق التي تتجاوز هذه المشكلة هي أن نربع هذه الانحرافات (فتصبح جميعها موجبة) ثم، ندني مجموع هذه المربعات.

وتسمى هذه بطريقة المربعات الصغرى للحصول على مقدري a و b . وتمثل المشكلة عندئذ في إيجاد الخط الذي يتوسط نقاط الانتشار حيث يجعل المجموع:

$$S = \sum (Y_t - \hat{Y}_t)^2 = \sum (Y_t - \hat{a} - \hat{b}X_t)^2$$

أقل ما يمكن.

وباستخدام حساب التفاضل. تعد هذه مشكلة مباشرة لتحديد قيمتي \hat{a} و \hat{b} اللتان تدنيان $E(Y_t - \hat{Y}_t)^2$. وينصح القارئ الذي لديه فكرة عن المبادئ الأساسية لحساب التفاضل أن يقوم بعمل الاشتقاقات المطلوبة التي هي معروضة أصلاً في ملحق هذا الفصل. وكل الذي نريدك أن تعرفه هو أننا عندما نحسب مقدري a و b باستخدام طريقة المربعات الصغرى، فإننا نحصل بالضبط على النتائج نفسها التي توصلنا إليها باستخدام طريقة المتغير المساعد. أي أن مقدرات طريقة المربعات الصغرى هي، أيضاً:

$$\hat{b} = \frac{\sum (X_t - \bar{X})(Y_t - \bar{Y})}{\sum (X_t - \bar{X})^2}$$

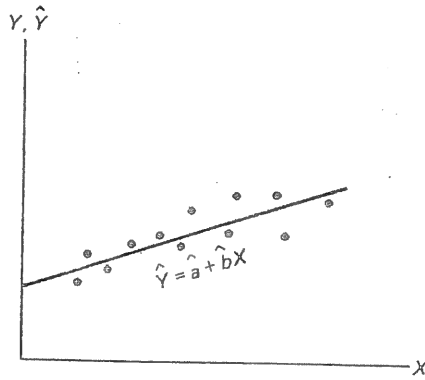
$$\hat{a} = \bar{Y} - \hat{b}\bar{X}$$

ويعد هذا أمراً مهماً لأنه عادة ماير القارئ في الأدب الاقتصادي على المعادلات التي أوضح المؤلف أنه توصل إليها بطريقة المربعات الصغرى، ومن الضروري أن يعرف أن هذه المعادلات متطابقة مع تلك المعادلات التي توصل إليها باستخدام إجراء التقدير الذي طور في هذا الفصل. ولقد حدث أن مثل هذا الإجراء يؤدي إلى تدنية مجموع مربعات انحرافات القيم المشاهدة عن القيم المحسوبة لـ Y .

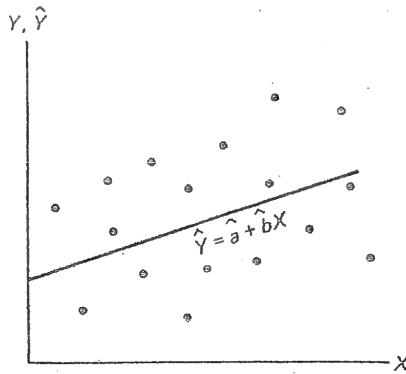
(٦-٢) قياس القوة التفسيرية لنموذج الانحدار

أصبح لدينا الآن طريقة لتقدير ما أسميناه بالعلاقة المتوسطة بين متغيرين، حيث تمكننا هذه الطريقة من الحصول على مقدرين للمعلمتين a و b في نموذج

الانحدار. غير أنه لا يوجد لدينا، حتى الآن، مقياس لدرجة قوة هذه العلاقة. على سبيل المثال، يلاحظ أن العلاقة بين X و Y في الشكلين رقم (٢-١٧) و (٢-١٨) واحدة، غير أن هاتين العلاقتين تختلفان في جانب مهم ألا وهو أن انتشار النقاط في الشكل الأول أقرب بكثير من الخط المستقيم منها بالشكل الثاني. وبمعنى آخر، لدينا تمثيل أدق "tighter fit" للنقاط المشاهدة على خط الانحدار.



شكل (٢-١٧)



شكل (٢-١٨)

ومن ثم، بالإضافة إلى وجود مقدرين لـ a و b من المهم أن تطور قياسا لهذا الجانب الآخر من العلاقة بين X و Y ، أي أننا نريد معرفة إلى أي مدى يعد نموذجنا جيدا. ونحن، عموما، نحاول أن نفسر التغير في القيم المشاهدة للمتغير Y باستخدام هذا النموذج. وإذا لم يكن لدينا نموذج فليس بوسعنا تفسير تحركات Y وأقصى ما يمكن عمله في هذه الحالة هو أن تأخذ \bar{Y} على أنها القيمة المتنبأ بها لـ Y بغض النظر عن قيمة X . والسؤال الآن هو ما إذا كان نموذجنا يمكنه أن يسمح لنا بعمل شيء أفضل من هذا، وإذا كان هذا هو الأمر، فبأي مقدار. ولهذا السبب، فسوف نقدم مقياسا للمقدرة التفسيرية لنموذج الانحدار الذي لدينا. أي أننا سوف نقدم مقياسا لمقدار التغير في Y الذي يمكن تفسيره بدلالة العلاقة الخطية المقدرة بين X و Y .

معامل التحديد The Coefficient of Determination

دعنا الآن نفحص شكل الانتشار (٢-١٩) والذي قدرنا له خط الانحدار الممثل بالمعادلة $\hat{Y} = \hat{a} + \hat{b}X$. ويتضح بالشكل متوسطي العينة \bar{X} و \bar{Y} للمتغيرين X و Y . وتوضح (2.45) إحدى خصائص المعادلة المقدرة والتي تتمثل في:

$$\bar{Y} = \hat{a} + \hat{b}\bar{X}, \quad (2.45)$$

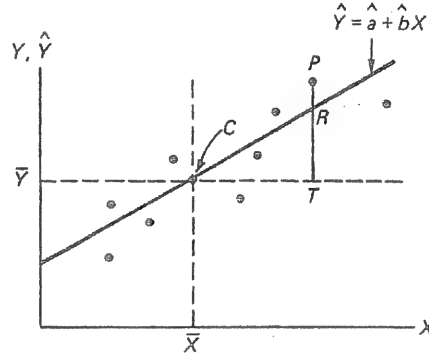
وتعني هذه المعادلة أن خط الانحدار يمر عبر النقطة (\bar{X}, \bar{Y}) التي هي النقطة "C" في الشكل رقم (٢-١٩). وبالنظر بعد ذلك إلى المشاهدة الممثلة بالنقطة P، نجد أن انحراف قيمة Y عند النقطة P عن قيمة متوسطها بالعينة \bar{Y} يساوي PT. كما سوف نلاحظ أن جزءا من انحراف Y عن وسطها \bar{Y} يمكن تفسيره بمعادلة الانحدار. وعلى وجه التحديد، فإن المعادلة المقدرة تفسر الجزء RT وتترك الجزء PR من الانحراف بدون تفسير. ويمكن التعبير عن هذه المسافات كمايلي:

* في الحقيقة، يتوافر لنا مقياسا لهذا الأمر، وذلك من المعالجة السابقة وهو تباين الخطأ العشوائي. فعلى سبيل المثال، يوضح شكل رقم (٢-١٧) أن هناك تباينا أقل للخطأ العشوائي بالمقارنة بالشكل رقم (٢-١٨).

PT = $Y_t - \bar{Y}$ = الانحراف الكلي لـ Y_t عن متوسط العينة

RT($\hat{Y}_t - \bar{Y}$) = انحراف Y_t عن \bar{Y} المفسر

PR($Y_t - \hat{Y}_t$) = انحراف Y_t عن \bar{Y} غير المفسر



شكل (٢-١٩)

ومع وجود هذه الخلفية، يمكننا الآن أن نشق مقياسا للمقدرة التفسيرية لمعادلة الانحدار. فلتذكر أولا من المعادلة (2.38) أن:

$$Y_t = \hat{Y}_t + \hat{u}_t \quad (2.38)$$

وبتجميع جانبي المعادلة (2.38) نحصل على:

$$\Sigma Y_t = \Sigma \hat{Y}_t + \Sigma \hat{u}_t \quad (2.104)$$

ولكن حيث إننا افترضنا مسبقا أن $\Sigma \hat{u}_t = 0$ ، فإن:

$$\Sigma Y_t = \Sigma \hat{Y}_t \quad (2.105)$$

وبقسمة طرفي المعادلة (2.105) على n نحصل على:

$$\bar{Y} = \bar{\hat{Y}} \quad (2.106)$$

وسوف تكون هذه النتائج مفيدة فيما بعد. دعنا نعود الآن إلى (2.38)، فبتربيع طرفيها، نحصل على:

$$Y_t^2 = \hat{Y}_t^2 + \hat{u}_t^2 + 2\hat{u}_t \hat{Y}_t \quad (2.107)$$

وبتجميع مشاهدات العينة، نحصل على:

$$\Sigma Y_t^2 = \Sigma \hat{Y}_t^2 + \Sigma \hat{u}_t^2 + 2\Sigma(\hat{u}_t \hat{Y}_t) \quad (2.108)$$

وبما أن:

$$\Sigma(\hat{u}_t \hat{Y}_t) = 0$$

ويتبع هذا بحكم أننا فرضنا الشرط

$$\Sigma(\hat{u}_t X_t) = 0 \quad \text{و} \quad \Sigma \hat{u}_t = 0$$

في طريقة التقدير. وطالما أن: $\hat{Y}_t = \hat{a} + \hat{b}X_t$ ، فإن:

$$\Sigma(\hat{u}_t \hat{Y}_t) = \hat{a}\Sigma \hat{u}_t + \hat{b}\Sigma(\hat{u}_t X_t) = 0$$

وهذا يعني أن الحد الأخير في المعادلة (2.108) يساوي الصفر، وبالتالي تصبح هذه المعادلة:

$$\Sigma Y_t^2 = \Sigma \hat{Y}_t^2 + \Sigma \hat{u}_t^2 \quad (2.109)$$

وبطرح $n\bar{Y}^2$ من المعادلة (2.109) نحصل على:

$$\Sigma Y_t^2 - n\bar{Y}^2 = (\Sigma \hat{Y}_t^2 - n\bar{Y}^2) + \Sigma \hat{u}_t^2 \quad (2.110)$$

وحيث إننا قد أوضحنا أن $\bar{Y} = \overline{\hat{Y}}$ ، فإنه يمكننا التعبير عن (2.110) بالصيغة التالية:

$$\Sigma(Y_t - \bar{Y})^2 = \Sigma(\hat{Y}_t - \bar{Y})^2 + \Sigma \hat{u}_t^2 \quad (2.111)$$

وتعد المعادلة (2.111) مفيدة جدا لخدمة أغراضنا فيما يتعلق بقياس المقدرة التفسيرية، ولذا، فإنه من المهم أن نفحص بعناية معنى كل حد من حدودها. ويلاحظ في هذا الصدد أن الجانب الأيسر من هذه المعادلة يعبر عن مجموع مربعات انحرافات Y_t عن متوسطها المقدر من العينة. ويعد هذا مقياسا للتغير في المتغير التابع الذي نبحث عن تفسير له من خلال معادلة الانحدار. وهذا يعني بديها أننا نريد أن يشرح نموذج الانحدار لدينا لماذا لا يبقى المتغير التابع Y_t ثابتا دائما. دعنا الآن نسمي الحد الأول على الجانب الأيسر من المعادلة (2.111) المجموع الكلي للمربعات Total Sum of Squares (TSS). ويأتي الدور الآن لفحص الجانب الأيمن من المعادلة

(2.111). وطالما أن $\hat{u}_i = (Y_i - \hat{Y}_i)$ ، فإن \hat{u}_i تشير إلى الجزء الذي لم يمكن تفسيره. أي أن \hat{u}_i تمثل انحراف القيمة المشاهدة للمتغير Y_i عن القيمة المحسوبة من معادلة الانحدار والتي تأخذ الصيغة التالية: $\hat{Y}_i = (\hat{a} + \hat{b}X_i)$ ، ويشير الحد الأخير في المعادلة (2.111) $\sum \hat{u}_i^2$ إلى مجموع مربعات الأخطاء، أي الجزء غير المفسر من التغير في Y_i . وسوف نطلق على هذا المجموع تسمية «مجموع مربعات الأخطاء» Error Sum of Squares (ESS). ويمثل الفرق بين TSS و ESS الحد $(\hat{Y}_i - \bar{Y})^2$ في المعادلة (2.111)، ومن الواضح أنه يعبر عن الجزء الذي يفسره نموذج الانحدار من المجموع الكلي للمربعات. وسوف نسمي هذا الجزء «مجموع مربعات الانحدار» Regression Sum of Squares (RSS) (أي مجموع المربعات المفسر بنموذج الانحدار). وتكون لدينا بذلك معادلة مقابلة للمعادلة (2.111) هي:

$$TSS = RSS + ESS \quad (2.112)$$

وقد يكون من المفيد أن نشرح هذه المعادلة باستخدام الشكل رقم (٢-١٤). فإذا أخذنا القيمة المشاهدة للمتغير Y_i الممثلة بالنقطة P، فإننا نجد أن PT تمثل انحراف Y_i عن متوسط العينة \bar{Y} بينما تمثل PR انحراف Y_i عن خط الانحدار، وبمعنى آخر، ذلك الجزء من انحراف Y_i عن \bar{Y} الذي لا يمكن تفسيره بدلالة خط الانحدار. أما عن المكون الآخر لـ PT الذي هو، على وجه التحديد، RT فإنه يمثل ذلك الجزء من Y_i الذي يفسره خط الانحدار. ويعني كل ما أوضحناه في المعادلة (2.111) أننا إذا حددنا ما يقابل المسافات PT، PR و RT لكل نقطة من نقاط العينة، فإن مجموع مربعات المسافات المقابلة لـ PT يساوي مجموع مربعات المسافات المقابلة لـ PR مضافا إليها مجموع مربعات المسافات المقابلة لـ RT. ونلخص ما سبق كما يلي:

$$TSS = (Y_i - \bar{Y})^2 = \text{المجموع الكلي للمربعات}$$

$$RSS = (\hat{Y}_i - \bar{Y})^2 = \text{مجموع مربعات الانحدار (المفسرة)}$$

$$ESS = \sum \hat{u}_i^2 = \text{مجموع مربعات الأخطاء (غير المفسرة)}$$

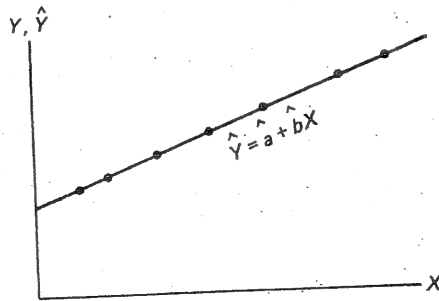
حيث: $TSS = RSS + ESS$. وطالما أن $ESS \geq 0$ و $RSS \geq 0$ ، فإنه يترتب على ذلك أن تكون $TSS \geq RSS$ و $TSS \geq ESS$. ولكي نقيس المقدرة التفسيرية لمعادلة الانحدار، فإننا في حاجة لمقياس يوضح النسبة التي يمكن لمعادلة الانحدار أن تفسرها من التغير في Y_t ، ويتمثل هذا المقياس في:

$$R^2 = \frac{RSS}{TSS} = 1 - \frac{ESS}{TSS}, \quad (2.113)$$

ويسمى R^2 معامل التحديد. ولو أن معادلة الانحدار تفسر كل التغير في Y_t (أي أن $\hat{Y}_t = Y_t$ لكل قيم t)، فإن $\hat{u}_t = 0$ ، ومن ثم، فإن $ESS = 0$. وفي هذه الحالة تكون $RSS = TSS$ ، وبالتالي فإن $R^2 = 1$. وطالما أن $Y_t = \hat{Y}_t = \hat{a} + \hat{b}X_t$ فإن Y_t سوف تكون على علاقة تامة مع X_t . ولهذا، فإن جميع نقاط شكل الانتشار بين Y_t و X_t سوف تقع على خط مستقيم (أي أن جميع قيم حد الخطأ سوف تساوي صفر). وتوضح هذه الحالة بالشكل (٢-٢٠).

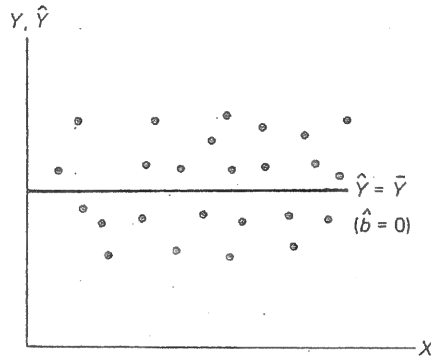
وعلى النقيض من ذلك، لو أن معادلة الانحدار لا تفسر أي جزء، فإن ESS تصل لحدها الأقصى، وعلى وجه التحديد فإن $ESS = TSS$ ومن ثم، فإن $RSS = 0$ ، $R^2 = 0$ ، فإن $\hat{Y}_t = \bar{Y}$ لكل قيم t . وهذا يتضمن أن $\hat{b} = 0$. وللتحقق من ذلك، دعنا نحل $\hat{a} = \bar{Y} - \hat{b}\bar{X}$ في $\hat{Y}_t = \hat{a} + \hat{b}X_t$ لكي نحصل على:

$$\hat{Y}_t = \bar{Y} + \hat{b}(X_t - \bar{X}) \quad (2.114)$$



شكل (٢-٢٠)

وحيث إن قيم X_i ليست كلها متساوية، أي أن $X_i \neq \bar{X}$ ، فإن هذا يعني أن كون $\hat{Y}_i = \bar{Y}$ يتضمن أن $\hat{b} = 0$. وفي هذه الحال، لا يعد النموذج ملائماً قطعياً حيث إن القيم المحسوبة لـ Y_i وبالتحديد \hat{Y}_i لا تعتمد البتة على قيم المتغير X_i ويوضح الشكل (٢١-٢) مثل هذه الحالة.



شكل (٢١-٢)

وفي الحالات العادية مثل تلك التي يوضحها شكلان (١٧-٢) و (١٨-٢) فإن معادلة الانحدار تفسر بعض التغير في Y وليس كله، ومن ثم، فإن R^2 تقع بين الصفر والواحد الصحيح. وكلما زادت النسبة التي تفسرها معادلة الانحدار من التغير في Y (أي كلما اقتربت نقاط الانتشار من خط الانحدار) اقتربت قيمة R^2 من الواحد الصحيح، وكلما كانت العلاقة بين X و Y أضعف كلما اقتربت قيمة R^2 من الصفر. ولعل هذا يعني أن R^2 توضح النسبة التي تفسرها المعادلة المقدرة من التغير في المتغير التابع. ومن ثم، فلو أن $R^2 = 0.63$ ، مثلاً، فإن هذا يعني أن العلاقة المقدرة تفسر ٦٣٪ من التغير في المتغير التابع.

$$R^2 = \hat{\rho}_{Y,\hat{Y}}^2$$

لقد قدمنا في القسم الأول من هذا الفصل مقياساً لمدى قوة الاقتران الخطي بين متغيرين، وأسميناه معامل الارتباط. عرف هذا المعامل بأنه يتمثل في نسبة

التغاير بين المتغيرين إلى حاصل ضرب انحرافيهما المعياريين . ويمكننا، أيضا أن نستخدم معامل الارتباط في قياس قوة العلاقة في معادلة الانحدار . فعلى سبيل المثال، يمكن لمعامل الارتباط بين Y_t و \hat{Y}_t أي $\rho_{Y, \hat{Y}}$ أن يقيس مدى اقتراب Y_t و \hat{Y}_t . ومن ثم، يمكن أن يؤخذ مقياسا لمدى مقدرة نموذج الانحدار على تفسير قيم Y_t .

وللأسف، فإن $\rho_{Y, \hat{Y}}$ سوف يكون، عموما، غير معلوم . ومن ثم، يتعين تقديره . واتساقا مع المعادلة (2.16) في القسم الأول من هذا الفصل، فإن مقدر $\rho_{Y, \hat{Y}}$ سوف يكون:

$$\hat{\rho}_{Y, \hat{Y}} = \frac{\Sigma(Y_t - \bar{Y})(\hat{Y}_t - \bar{Y})}{\sqrt{\Sigma(Y_t - \bar{Y})^2 \Sigma(\hat{Y}_t - \bar{Y})^2}} \quad (2.115)$$

طالما أن $\overline{\hat{Y}} = \bar{Y}$.

وسوف نثبت الآن أن $\hat{\rho}_{Y, \hat{Y}}^2 = R^2$ أي أن R^2 هي، ببساطة، مربع مقدر معامل الارتباط بين Y_t و \hat{Y}_t . ولذا لا يعد R^2 و $\hat{\rho}_{Y, \hat{Y}}$ مقياسان بديلين لقوة العلاقة بين Y و \hat{Y} .

دعنا الآن نفحص بسط المعادلة (2.115)، فنحن نتذكر من الملحق A بالفصل الأول أنه، لأي متغيرين، وليكونا Z_{1t} و Z_{2t} ، بمتوسطي عينة \bar{Z}_1 و \bar{Z}_2 نجد أن:

$$\Sigma(Z_{1t} - \bar{Z}_1)(Z_{2t} - \bar{Z}_2) = \Sigma(Z_{1t} - \bar{Z}_1)Z_{2t}$$

حيث يمكن تبسيط بسط المعادلة (2.115) فيما يلي:

$$\Sigma(\hat{Y}_t - \bar{Y})Y_t$$

وبما أننا نعرف أن $Y_t = \hat{Y}_t + \hat{u}_t$ فإننا نحصل على:

$$\Sigma(\hat{Y}_t - \bar{Y})Y_t = \Sigma(\hat{Y}_t - \bar{Y})(\hat{Y}_t - \hat{u}_t) = \Sigma(\hat{Y}_t - \bar{Y})\hat{Y}_t$$

طالما أن:

$$\Sigma(\hat{Y}_t \hat{u}_t) = 0 \text{ و } \Sigma(\hat{u}_t \bar{Y}) = \bar{Y} \Sigma \hat{u}_t = 0$$

وأخيراً، يمكن وضع البسط في الصيغة التالية:

$$\Sigma(\hat{Y}_t - \bar{Y})Y_t = \Sigma(\hat{Y}_t - \bar{Y})(\hat{Y}_t - \bar{Y}) = \Sigma(\hat{Y}_t - \bar{Y})^2 = RSS.$$

وبملاحظة أن مقام $\hat{\rho}_{Y, \hat{Y}}$ هو:

$$\sqrt{\Sigma(Y_t - \bar{Y})^2 \Sigma(\hat{Y}_t - \bar{Y})^2} = \sqrt{(TSS)(RSS)}$$

فإن:

$$\rho_{Y, \hat{Y}} = \frac{RSS}{\sqrt{(RSS)(TSS)}} = \frac{\sqrt{RSS}}{\sqrt{TSS}} \quad (2.116)$$

ومن المعادلة (2.116) نلاحظ أن $R^2 = \hat{\rho}_{Y, \hat{Y}}^2$

مثال

قد يكون من المفيد، لغرض التوضيح، أن نعود مرة أخرى لدالة الاستهلاك التي قدرناها سابقاً في هذا الفصل ونوجد قيمة R^2 . ولتبسيط الحسابات، دعنا نستخدم الصيغة التالية:

$$R^2 = \frac{\Sigma(\hat{Y}_t - \bar{Y})^2}{\Sigma(Y_t - \bar{Y})^2} = \frac{\Sigma(\hat{Y}_t - \bar{Y})\hat{Y}_t}{\Sigma(Y_t - \bar{Y})\hat{Y}_t} = \frac{\Sigma\hat{Y}_t^2 - n\bar{Y}^2}{\Sigma Y_t^2 - n\bar{Y}^2}, \quad (2.117)$$

حيث إن $\bar{Y} = \bar{\hat{Y}}$ كما أثبتنا سابقاً.

والخطوة التالية لحساب R^2 هي أن نستخدم معادلة الانحدار المقدرة لحساب

القيم المقدرة لـ Y (أي \hat{Y}_t) ويوضح الجدول رقم (٢-٨) هذه الحسابات.

جدول رقم (٢-٨)^أ

$$\hat{C} = 13 + 0.89 Y_d$$

C	Y_d	C2	\hat{C}	\hat{C}^2
325	350	105,625	325	105,625
335	346	112,225	337	113,569
355	385	126,025	356	126,736
375	405	140,625	373	139,129
401	438	160,801	403	162,409
433	473	187,489	434	188,356
466	512	217,156	469	219,961
492	547	242,064	500	250,000
537	590	288,369	538	289,444
576	630	331,776	574	329,476

$$\Sigma \hat{C}^2 = 1,912,155$$

$$\Sigma \hat{C}^2 = 1,924,705$$

$$\bar{C}^2 = 184,900$$

$$n\bar{C}^2 = 1,849,000$$

$$R^2 = \frac{\Sigma \hat{C}_i^2 - n\bar{C}^2}{\Sigma C_i^2 - n\bar{C}^2} = 0.99$$

(a) لأن قيم المعاملات في معادلة الانحدار مقربة إلى رقمين عشريين، فقط، فإن قيمة R^2 المحسوبة من الأرقام أكبر من الواحد بشئ بسيط نتيجة أخطاء التقريب. ولو أن المعاملات حسبت باستخدام كل الأرقام

العشرية الموجودة فإن قيمة R^2 سوف تكون 0.99

ومن الواضح أن قيمة R^2 لدالة الاستهلاك المقدرة هي 0.99، الأمر الذي يشير إلى وجود اقتران قوي جدا بين C و Y_d . وهي تعني أن معادلة الانحدار المقدرة تفسر ٩٩٪ من التغير في C ، وأن ١٪، فقط، يبقى غير مفسر. وهذا يؤكد النتيجة التي توصلنا إليها سابقا من خلال النظر إلى شكل الانتشار وخط الانحدار بالشكل (٢-١٤).

(٢-٧) توضيح: تقدير دالة تكلفة

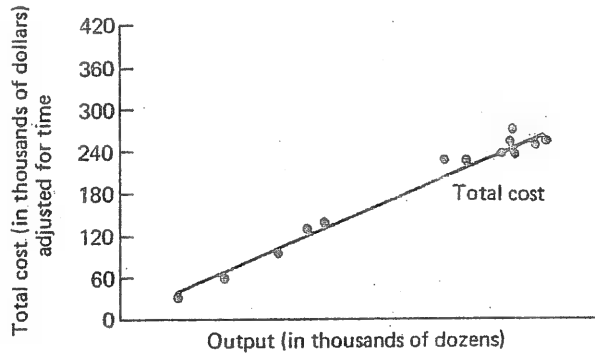
ونختتم هذه المقدمة عن نموذج الانحدار البسيط بدراسة تطبيقية من الأدب

الاقتصادي. ومن النقاط التي تركز نظرية الاقتصاد الجزئي للمنشأة على توضيحها دالة تكاليف المنشأة. فمعظم المراجع تقدم مناقشة مطولة على وجه الخصوص للعلاقة بين التكاليف ومستوى إنتاج المنشأة. ويمكن صياغة دالة التكاليف في الصورة العامة التالية:

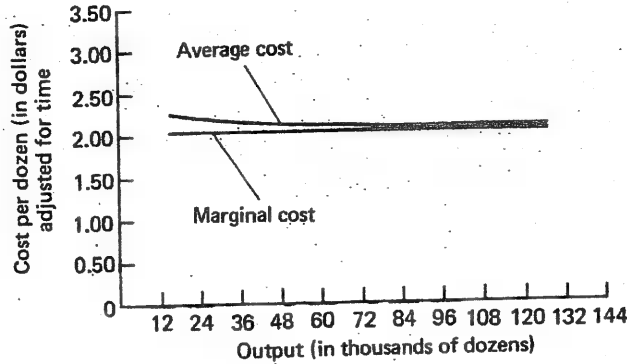
$$C = f(Q),$$

حيث C التكاليف و Q حجم ناتج المنشأة وللحصول على مزيد من المعرفة عن شكل هذه العلاقة، فإن بعض الاقتصاديين استخدموا تحليل الانحدار في تقدير دوال التكلفة من بيانات فعلية عن التكاليف والناتج.

ومن الدراسات المبكرة والرائدة في هذا المجال دراسة *Joel Dean عن دالة التكاليف لمصنع جوارب حريرية. فلقد قام Dean بجمع بيانات شهرية عن إنتاج المصنع وتكاليف. ويعرض الشكل (٢-٢٢) هذه البيانات في شكل انتشار. ويلاحظ أن النقاط تقترب في انتشارها إقترابا كبيرا من الخط المستقيم بالشكل. وهذا يعني أن الصيغة الخطية سوف تصف دالة تكاليف المصنع جيدا. وباستخدام نموذج انحدار المتغيرين نجد أن:



شكل رقم (٢-٢٢)



شكل رقم (٢-٢٣)

$$C_i = a + bQ_i + u_i \quad (2.118)$$

حيث: C = التكاليف الكلية شهريا مقاسة بالآلاف دولار،

Q = الكمية المنتجة شهريا مقاسة بالآلاف دسته من أزواج الجوارب،

u = خطأ عشوائي.*

وتتضمن الدالة (2.118) بعض النقاط المهمة لطبيعة تكاليف المصنع. فحيث إن هذه الدالة هي دالة تكاليف كلية فإن التكلفة الحدية المتوقعة والمصاحبة لكل وحدة مضافة شهريا من الناتج تساوي b وحدة نقدية. ووفقا لوحدة القياس المستخدم، فإن هذا يعني أنه إذا زاد الناتج الشهري بمقدار ألف دسته، فإن التكاليف الشهرية سوف تزداد بمقدار b ألف دولار. ويتضمن هذا بدوره أن كل زيادة في الناتج الشهري بمقدار دسته واحدة تؤدي إلى زيادة التكاليف الشهرية بمقدار b دولار. وتوضح المعلمة a أن التكاليف الشهرية سوف تساوي a ألف دولار إذا كان مستوى الناتج الشهري مساويا صفر. أي أنها تمثل التكاليف الثابتة الشهرية للمنشأة. ومن الملاحظ أنه إذا كانت a موجبة فإن متوسط التكلفة الثابتة الشهرية للمصنع سوف تتناقص مع تزايد الناتج، طالما أن التكاليف الثابتة تتوزع على وحدات إنتاج أكثر.

* وهذا كمثال لتوضيح القياسات بقيمة $C = 27$ تقابل $27,000$ \$، بقيمة $Q = 17$ تقابل $17,000$ دسته أو $(17,000 \times 12) = 204,000$ زوج جوارب

وقدر Dean معادلة الانحدار (2.118) باستخدام طريقة المربعات الصغرى التي تكافئ طريقة المتغير المساعد التي عرضت في هذا الفصل، فجاءت على النحو التالي:

$$C = 2.936 + 2.00Q \quad (2.119)$$

$$R^2 = 0.95$$

وكما يوضح شكل الانتشار، فإن توفيق الخط للنقاط جيد، ويبلغ معامل التحديد 0.95 وهو ما يعني أن المعادلة المقدرة (2.119) تفسر ٩٥٪ من التغير المشاهد في التكاليف الكلية. وبالإضافة إلى ذلك، فإن القيم المقدرة للمعاملات تمدنا ببعض المعلومات المحددة عن التكاليف. فهي توضح أن التكاليف الثابتة للمصنع تبلغ ٢٩٣٦ \$ شهريا وأن التكلفة الحدية لدسته أزواج الجوارب تساوي ٢ \$. ويوضح الشكل (٢-٢٣) دالتي التكلفة الحدية والمتوسطة. وتعد مثل هذه المعلومات مهمة ليس، فقط، للاقتصادى الذي يدرس علاقات التكلفة، وإنما أيضا لإدارة المنشأة.

ملحق: إثباتات لثلاث نتائج

لقد أشرنا في المتن إلى ثلاث نتائج، إحداها تتعلق بتباين مجموع المتغيرات العشوائية غير المرتبطة، وأخرى تتعلق بخاصية التباين الأدنى لمقدرات طريقة المتغير المساعد، والثالثة تتعلق بخاصية المربعات الصغرى للمقدرات، وسوف نقدم هنا اثباتات لهذه النتائج.

تباين مجموع المتغيرات العشوائية

دعنا نفترض أن X_1, X_2, \dots, X_n متغيرات عشوائية، وأن a_1, a_2, \dots, a_n ثوابت، وأن متوسطات هذه المتغيرات وتبايناتها على التوالي هي $\mu_1, \mu_2, \dots, \mu_n$ و $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$. ودعنا نعرف Y كمايلي:

$$Y = a_0 + a_1X_1 + \dots + a_nX_n \quad (2A.1)$$

ويتعين ملاحظة أنه، طالما أن Y مكونة من توليفة خطية من مجموعة متغيرات عشوائية، فإن Y نفسها تعد متغيراً عشوائياً. وسوف نوضح مايلي:

الأول: إذا كانت X_1, X_2, \dots, X_n متغيرات عشوائية غير مرتبطة، وبجعل σ_Y^2 تشير إلى تباين Y عندئذ:

$$\sigma_Y^2 = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \dots + a_n^2 \sigma_n^2 \quad (2A.2)$$

وتعني المعادلة (2A.2) أنه إذا كانت قيم المتغيرات X 's غير مرتبطة خطيا مع بعضها بعضا، فإن تباين Y يساوي مجموع حاصل ضرب تباينات المتغيرات X ، كل مضروباً بمربع معامل المرتبط به.

ولكي نثبت النتيجة الأولى، يتعين أولاً ملاحظة أن متوسط Y ، أي $E(Y) = \mu_Y$ ، هو، ببساطة، متوسط الجانب الأيمن من المعادلة (2A.1):

$$\begin{aligned} \mu_Y &= E(a_0) + E(a_1 X_1) + \dots + E(a_n X_n) \\ &= a_0 + a_1 \mu_1 + a_2 \mu_2 + \dots + a_n \mu_n \end{aligned} \quad (2A.3)$$

وحيث إن تباين Y يساوي $E(Y - \mu_Y)^2$ بالتعريف، وباستخدام كل من (2A.3) و (2A.1) فإننا نحصل على:

$$\sigma_Y^2 = E(Y - \mu_Y)^2 = E[a_1(X_1 - \mu_1) + a_2(X_2 - \mu_2) + \dots + a_n(X_n - \mu_n)]^2 \quad (2A.4)$$

وبفك المعادلة (2A.4)، نحصل على نوعين من الحدود، النوع الأول حدود مربعة والنوع الثاني مجموعة من حواصل ضرب تقاطعية:

$$\begin{aligned} \sigma_Y^2 &= E[a_1^2(X_1 - \mu_1)^2 + a_2^2(X_2 - \mu_2)^2 + \dots + a_n^2(X_n - \mu_n)^2 \\ &\quad + \dots + 2a_3a_4(X_3 - \mu_3)(X_4 - \mu_4) + \dots] \end{aligned} \quad (2A.5)$$

ويلاحظ أن الحد الأخير في (2A.5) هو حاصل ضرب مجموعة من العناصر. وقد يوجد هناك عدد كبير من هذه العناصر، وفقاً لمدى كبر n . والنقطة الأساسية هنا هي، طالما أن المتغيرات X 's غير مرتبطة افتراضاً فإن القيمة المتوقعة لحاصل ضرب كل عنصرين من عناصر الحد الأخير سوف تساوي صفراً. وللتوضيح، فإن هذا يعني لـ (2A.5) أن:

$$\begin{aligned}
 E\left[2a_3a_4(X_3 - \mu_3)(X_4 - \mu_4)\right] &= 2a_3a_4E(X_3 - \mu_3)(X_4 - \mu_4) \\
 &= 2a_3a_4\text{cov}(X_3, X_4) = 0
 \end{aligned}
 \quad (2A.6)$$

حيث $\text{cov}(X_3, X_4)$ هي التغاير بين X_3 و X_4 ، وبما أن القيمة المتوقعة لحاصل ضرب كل عنصرين تقاطعين تساوي الصفر، فإن المعادلة (2A.5) تصبح:

$$\begin{aligned}
 \sigma_Y^2 &= E\left[a_1^2(X_1 - \mu_1)^2 + a_2^2(X_2 - \mu_2)^2 + \dots + a_n^2(X_n - \mu_n)^2\right] \\
 &= a_1^2E(X_1 - \mu_1)^2 + a_2^2E(X_2 - \mu_2)^2 + \dots + a_n^2E(X_n - \mu_n)^2 \\
 &= a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2
 \end{aligned}
 \quad (2A.7)$$

المقدرات ذات أصغر تباين لكل من a و b *

لقد أوضحنا في المتن أن طريقة المتغير المساعد يتولد عنها مقدرات \hat{a} و \hat{b} غير متحيزة لمعاملات نموذج الانحدار. وسوف نثبت الآن أن مقدري طريقة المتغير المساعد \hat{a} و \hat{b} يتصفان بكون تباينهما المشروط هو الأقل من بين مجموعة المقدرات الخطية غير المتحيزة للمعلمتين a ، b . ويعرف هذا في الأدب القياسي بنظرية جاوس - ماركوف Gauss-Markov theorem.

ويلاحظ في البداية، من المتن، أن:

$$\begin{aligned}
 \hat{b} &= \frac{\sum (X_t - \bar{X})Y_t}{\sum (X_t - \bar{X})^2} = \frac{\sum w_t Y_t}{A} \\
 &= \sum Q_t Y_t
 \end{aligned}
 \quad (2A.8)$$

حيث $w_t = (X_t - \bar{X})$ و $A = \sum (X_t - \bar{X})^2$. وبالمقابل فقد ورد في المتن أن:

* يعتمد هذا القسم على : J. Johnston, *Econometric Methods*, 2nd ed. (New York: McGraw-Hill, 1972), pp. 18-23.

$$\begin{aligned}\hat{a} &= \bar{Y} - \hat{b}\bar{X} \\ &= \Sigma \gamma_i Y_i\end{aligned}\quad (2A.9)$$

حيث $\gamma_i = (1/n) - \bar{X}(w_i/A)$. ويلاحظ، كما في المتن، أن كلا من Q_i و γ_i يعتمد على قيم X_i فقط. ولذا، فلو أن قيم X_i كانت معطاة يمكننا اعتبار قيم Q_i و γ_i ثوابت.

ونريد الآن أن نحدد خاصيتين للأوزان Q_i . فمن الملاحظ أولاً أن:

$$\Sigma Q_i = \Sigma \left(\frac{X_i - \bar{X}}{A} \right) = \frac{1}{A} \Sigma (X_i - \bar{X}) = 0 \quad (2A.10)$$

وثانياً:

$$\Sigma (Q_i X_i) = \frac{1}{A} \Sigma (X_i - \bar{X}) X_i = \frac{1}{A} \Sigma (X_i - \bar{X})^2 = \left(\frac{1}{A} \right) A = 1, \quad (2A.11)$$

حيث، كما نذكر، فإن:

$$\Sigma (X_i - \bar{X}) X_i = \Sigma (X_i - \bar{X}) (X_i - \bar{X}) = \Sigma (X_i - \bar{X})^2 = A$$

ولقد أوضحنا في المتن أن:

$$\sigma_b^2 = \frac{\sigma_u^2}{\Sigma (X_i - \bar{X})^2} \quad (2A.12)$$

ويمكن التعبير عن (2A.12) على النحو التالي:

$$\sigma_b^2 = \sigma_u^2 \Sigma Q_i^2 \quad (2A.13)$$

حيث:

$$\Sigma Q_i^2 = \frac{\Sigma w_i^2}{A^2} = \frac{\Sigma (X_i - \bar{X})^2}{A^2} = \frac{A}{A^2} = \frac{1}{A} = \frac{1}{\Sigma (X_i - \bar{X})^2}$$

دعنا نفترض الآن أن \hat{b}^* هي أي مقدر خطي آخر للمعلمة b ، فعليه:

$$\begin{aligned}\hat{b}^* &= \Sigma(Q_t + v_t)Y_t \\ &= \hat{b} + \Sigma v_t Y_t\end{aligned}\quad (2A.14)$$

حيث v_t (مثل Q_t) دالة في X_t ولكن ليس في Y_t . وتقرر المعادلة (2A.14)، ببساطة، أن \hat{b}^* تساوي \hat{b} مضافا إليها قيمة أخرى تمثل الفرق بينهما. ومع الأخذ في الاعتبار أن $\Sigma Q_t = 0$ و $\Sigma(Q_t X_t) = 1$ ، فإنه يتبع ذلك أن تكون:

$$\begin{aligned}\hat{b}^* &= \Sigma(Q_t + v_t)Y_t = \Sigma(Q_t + v_t)(a + bX_t + u_t) \\ &= a\Sigma(Q_t + v_t) + b\Sigma(Q_t + v_t)X_t + \Sigma(Q_t + v_t)u_t \\ &= a\Sigma v_t + b + b\Sigma(v_t X_t) + \Sigma(Q_t + v_t)u_t\end{aligned}\quad (2A.15)$$

ومع تذكر أن قيم Q_t و v_t تعتمد فقط، على قيم X_t ، ولذا، فإنها قيم ثابتة، فإن القيمة المتوقعة لـ b تصبح:

$$E(\hat{b}^*) = a\Sigma v_t + b + b\Sigma(v_t X_t) \quad (2A.16)$$

حيث إن $E(u_t) = 0$.

ولو أن \hat{b}^* مقدر غير متحيز للمعلمة b ، فلا بد أن يتحقق الشرط التالي:

$$E(\hat{b}^*) = b.$$

ولكي يكون \hat{b}^* غير متحيز، فإن هذا يتضمن من المعادلة (2A.16)، أن تكون:

$$\Sigma v_t = 0 \text{ و } \Sigma(v_t X_t) = 0 \quad (2A.17)$$

وبهذه المعلومات، يمكننا إعادة كتابة الصيغة الأخيرة للمعادلة (2A.15) كما يلي:

$$\begin{aligned}\hat{b}^* &= b + \Sigma(Q_t + v_t)u_t \\ &= b + (Q_1 + v_1)u_1 + (Q_2 + v_2)u_2 + \dots + (Q_n + v_n)u_n\end{aligned}\quad (2A.18)$$

ويمكن الآن أن نستخدم النتيجة الأولى التي سبقت الإشارة إليها بالقسم الأول من هذا الملحق للحصول على صيغة لتباين \hat{b}^* . فطالما أن \hat{b}^* توليفة خطية من قيم u_t ، وبما أن قيم u_t s غير مرتبطة لأنها مستقلة، فإن تباين \hat{b}^* يساوي:

$$\begin{aligned}
\sigma_{\hat{b}}^2 &= (Q_1 + v_1)^2 \sigma_u^2 + \dots + (Q_n + v_n)^2 \sigma_u^2 \\
&= \sigma_u^2 (Q_1 + v_1)^2 \\
&= \sigma_u^2 [\Sigma Q_i^2 + \Sigma v_i^2 + 2\Sigma(Q_i v_i)] \\
&= \sigma_u^2 [\Sigma Q_i^2 + \Sigma v_i^2]
\end{aligned} \tag{2A.19}$$

بحكم أن:

$$2\Sigma(Q_i v_i) = \frac{2\Sigma(X_i - \bar{X})v_i}{A} = \frac{2}{A} [\Sigma(X_i v_i) - \bar{X}\Sigma v_i] = 0 \tag{2A.20}$$

ومن (2A.13) و (2A.19) يمكن أن نرى أن:

$$\begin{aligned}
\sigma_{\hat{b}}^2 &= \sigma_u^2 \Sigma Q_i^2 + \sigma_u^2 \Sigma v_i^2 \\
&= \sigma_b^2 + \sigma_u^2 \Sigma v_i^2
\end{aligned} \tag{2A.21}$$

وحيث إنه من الواضح أن الحد الأخير في (2A.21)، سيكون موجبا إذا كانت بعض قيم $v_i \neq 0$ ، يمكن أن نخلص إلى:

$$\sigma_{\hat{b}}^2 \geq \sigma_b^2 \tag{2A.22}$$

ومن الملاحظ أن $\sigma_{\hat{b}}^2$ سوف يساوي σ_b^2 إذا كانت $\Sigma v_i = 0$ ، فقط، وهذا لا يحدث إلا إذا كانت جميع قيم $v_i = 0$ ، وعندئذ، فإن $\hat{b}^* \equiv \hat{b}$. ولقد أوضحنا بذلك أن أي مقدر خطي آخر غير متحيز لـ b سوف يكون تباينه المشروط أكبر من تباين المقدر الخطي لطريقة المتغير المساعد. وباستخدام نفس الإجراء يمكن إثبات النتيجة نفسها للمقدر الخطي الخاص بالمعلمة a لطريقة المتغير المساعد. وسوف نترك هذا التمرين للقارئ المهتم.

خاصية أصغر المربعات لكل من \hat{a} و \hat{b}

كما هو موضح في المتن فإن مقدرات المربعات الصغرى تم اشتقاقها عن

طريق تصغير المجموع التالي لـ \hat{a} و \hat{b} :

$$S = \Sigma(Y_i - \hat{Y}_i)^2 = \Sigma(Y_i - \hat{a} - \hat{b}X_i)^2 \quad (2A.23)$$

وبمفاضلة المعادلة (2A.23) جزئيا لكل من \hat{a} و \hat{b} ومساواة النتيجة بالصفر نحصل على :

$$\begin{aligned} \frac{\partial S}{\partial \hat{a}} &= 2\Sigma(Y_i - \hat{a} - \hat{b}X_i)(-1) = 0 \\ \frac{\partial S}{\partial \hat{b}} &= 2\Sigma(Y_i - \hat{a} - \hat{b}X_i)(-X_i) = 0 \end{aligned} \quad (2A.24)$$

وبضرب المعادلتين (2A.24) في $(-\frac{1}{2})$ مع التبسيط ، نحصل على :

$$\Sigma Y_i = n\hat{a} + \hat{b}\Sigma X_i \quad \text{or} \quad \bar{Y} = \hat{a} + \hat{b}\bar{X} \quad (2A.25)$$

و

$$\Sigma(Y_i X_i) = \hat{a}\Sigma X_i + \hat{b}\Sigma X_i^2 \quad (2A.26)$$

وحيث إن (2A.25) متطابقة مع المعادلة الطبيعية الأولى (2.45)، والمعادلة (2A.26)، بعد قسمتها على n ، متطابقة مع المعادلة الطبيعية الثانية (2.50) كما وردت في المتن، فإننا نستخلص من ذلك أن مقدرات طريقة المربعات الصغرى متطابقة مع مقدرات طريقة المتغير المساعد.

أسئلة

١ - أثبت أنه يمكن التعبير عن مقدر الحد الثابت \hat{a} على النحو :

$$\hat{a} = \Sigma\left(\frac{1}{n} - \bar{X}W_i\right)Y_i$$

حيث :

$$W_i = \frac{(X_i - \bar{X})}{\Sigma(X_i - \bar{X})^2}$$

- ٢ - اعتبر نموذج الانحدار التالي $Y_i = a + bX_i + u_i$ ، حيث تأخذ المشاهدات حول X_i و Y_i القيم التالية:

X_i	Y_i
4	8
2	6
3	5
1	7
2	4

قدر كلا من a ، b و σ_u^2 .

- ٣ - اقترح أحد محللي الاستهلاك أن دالة الاستهلاك $C_i = a + bY_i$ لافائدة منها، لأن النقاط (C_i و Y_i) في شكل الانتشار لاتقع جميعها علي خط مستقيم. كما اقترح، أيضا، أنه أحيانا، تزداد Y_i في الوقت الذي تنخفض فيه C_i . واستنتج من ذلك أن C_i ليست دالة في Y_i . قوم هذه العبارة.

- ٤ - دع $Y = 5 - 3X$. اثبت أن معامل الارتباط $\rho_{X,Y} = \sigma_{X,Y} / \sigma_X \sigma_Y$ يكون مساويا لـ (١-).

- ٥ - دع قيم تباينات المتغيرات X_1 ، X_2 و X_3 هي 1، 3 و 5 على الترتيب. افترض أن هذه المتغيرات مستقلة. دع $Y = 13 - 2X_1 + 3X_2 - 10X_3$. أوجد تباين Y .

- ٦ - يقال إن الأسر متوسطة الدخل ومرتفعته تترك المدن وتذهب للإقامة في الضواحي لأن معدلات الضرائب داخل المدن أعلى من معدلاتها في الضواحي القريبة منها. افترض أنه يتوافر لنا بيانات حول عدد من المدن في سنة معينة. كون هذه الفرضية في شكل نموذج انحدار. (هناك أكثر من طريقة واحدة لعمل ذلك!).

- ٧ - اعتبر النموذج التالي:

$$Y_i = a_1 + b_1 X_i + u_i \quad (1)$$

حيث إن الخطأ u_i يعتمد على X_i على النحو التالي:

$$u_i = a_2 + b_2 X_i + \varepsilon_i \quad (2)$$

حيث إن ε_i هو خطأ عشوائي مستقل عن X_i ويحقق، أيضا افتراضاتنا المعتادة كافة. افترض أن $b_2 > 0$. بين أن b_1 في (١) يقلل تأثير X_i على Y_i .

٨ - افترض في السؤال السابع أنه تم إحلال معادلة (2) محل المعادلة:

$$u_i = a_2 + b_2 X_i^2 + \varepsilon_i$$

هل سيستهك في هذه الحال أي من افتراضاتنا المعتادة لنموذج الانحدار الثاني الذي يربط Y_i و X_i ؟

٩ - كون نموذج انحدار يصف العلاقة بين عمر الطفل وطوله. ناقش هل يعاني هذا النموذج بعض أوجه القصور أم لا؟
١٠ - اعتبر النموذج التالي:

$$Y_i = a + bX_i + u_i$$

حيث يوجد لدينا أخطاء في القياس نتيجة عدم ملاحظتنا لـ X_i مباشرة. بدلا من ذلك، افترض أننا لاحظنا:

$$X_i^m = X_i + \varepsilon_i$$

حيث ε_i هو خطأ عشوائي مستقل عن X_i ، وله قيمة متوسطة صفرية. ويحقق افتراضاتنا المعتادة كافة. إضافة إلى ذلك افترض أن ε_i و u_i مستقلان. ويتضمن هذا استقلال X_i^m و u_i .

(أ) - كون نموذج الانحدار الذي يربط Y_i بـ X_i^m .

(ب) - هل تم انتهاك أي من افتراضاتنا المعتادة في هذا النموذج؟

الفصل الثامن

تطبيقات نموذج الانحدار

أوضحنا في الفصل السابق نموذج الانحدار الأساسي ذي المتغيرين، وأوجدنا طريقة تقدير معلماته. والآن يمكننا، عن طريق استخدام عينة من القيم المشاهدة للمتغيرين المرتبطين، إيجاد مقدرات لكل من a و b في معادلة الانحدار، وإيجاد كذلك تباينات هذه المقدرات. كما يمكننا أيضا، قياس قوة العلاقة بين المتغيرين عن طريق حساب نسبة التغير في المتغير التابع التي يمكن أن تفسرها معادلة الانحدار المقدرة. وسنبين في هذا الفصل كيف يستخدم الاقتصاديون هذه الطرق لاختبار الفرضيات حول السلوك الاقتصادي لعمل التوقعات.

(١-٣) اختبار الفرضيات وفترات الثقة : مقدمة

وضعنا في الفصل الثاني فرضية أن الإنفاق الاستهلاكي يعتمد على مستوى الدخل المتاح، وقدرنا الشكل الخطي لهذه العلاقة، ووجدنا باستخدام السلاسل الزمنية للبيانات الكلية للولايات المتحدة الأمريكية أن القيمة المقدرة للميل الحدي للاستهلاك موجبة وتنحصر بين الصفر والواحد الصحيح، كما أن تقديرنا لـ a كان موجبا أيضا. وأشارنا إلى أن هذه النتائج تتفق مع النظرية الكينزية المعروفة بدالة الاستهلاك.

ولكن، هل يمكن الركون إلى هذه النتائج ؟ على سبيل المثال، ما مدى ثقتنا في كون المعلمة a موجبة في الحقيقة ؟ مثلا، لو كان تقديرنا لـ a هو 0.001، فهل يكون هذا مقنعا بالقول إن a موجبة وليست صفرا.

وبالمقابل، افترض أننا، وبناء على المعلومات المسبقة والمتاحة، نعتقد أن $b = .75$ ، فهل يكون تقديرنا لـ b ، تحديداً 0.89، غير متنسق مع الافتراض المسبق بأن $b = .75$ ؟ النقطة الأساسية هنا هي أن التفاوت بين تقديرنا لـ b ، 0.89، وبين القيمة المفترضة لها $b = .75$ نشأ من حجم العينة التي استخدمناها في التقدير. وقياساً على ذلك إذا كان طول الأفراد في الولايات المتحدة "10' 5" مثلاً، فلا ينبغي لنا أن نتوقع أن الطول المتوسط لمجموعة عشوائية من ثلاثة أفراد (مثلاً) يكون "10' 5" بالضبط. هذه هي مشاكل اختبار الفرضيات. نهتم، أساساً، في هذه المشاكل بمعرفة هل هناك تناسق بين تقديراتنا للمعلمات والفرضيات المسبقة أم لا. وهناك مشكلة وثيقة الصلة باختبارات الفرضيات هي المرتبطة بفترات الثقة. فتقدير معين لـ b ، تحديداً 0.89. ويدعي هذا تقدير النقطة point estimate. فإذا ما اضطررنا لتفسير هذا التقدير، فيحتمل أن نقول أن الميل الحدي للاستهلاك «حوالي» 0.89 أي يحتمل ألا نتوقع أن تكون قيمة b مساوية 0.89 بالضبط.

لذلك نرغب أن نبني فترة ثقة حول تقديرنا بالنقطة التي نشعر، بدرجة من الثقة، أنها تحتوي على قيمة b . وتؤدي نظرية فترات الثقة التي سوف نطورها أدناه هذه المهمة. إنها تمكننا من استخدام تقدير النقطة لإيجاد التقدير بالفترة، ومثل هذه الفترة مدى من القيم تؤدي، بسبب طريقة بناؤها، إلى توقع أن قيمة المعلمة موضع الاهتمام مشمولة ضمن هذه الفترة، عند مستوى معين من الثقة، وعلى سبيل المثال، يمكن أن تظهر عبارة فترة الثقة على النحو التالي: باحتمال قدرة 0.95، تحتوي الفترة $(\hat{b} \pm 0.07)$ على قيمة المعلمة b . سوف نبين، أخيراً، أن الطريقة الوحيدة التي نحصل بها على ثقة أو تأكيد أكبر بأن فترتنا تحتوي على b ، مع ثبات العوامل الأخرى، تكون من خلال توسيع تلك الفترة. فإذا كان لدينا 95% مستوى ثقة بأن b تقع في حدود 0.07 من b ، ولكننا نحتاج إلى فترة تشمل على b باحتمال قدره 99%، فإنه ينبغي أن تكون الفترة الجديدة أوسع من $(\hat{b} - 0.07, \hat{b} + 0.07)$ على سبيل المثال، يمكن أن تكون فترة الثقة 99% هي $(\hat{b} - 0.10, \hat{b} + 0.10)$.

وترتبط مشاكل اختبار الفرضيات وفترات الثقة ببعضها بعضا على النحو التالي : افترض ، كما ذكرنا من قبل ، أننا نرغب في اختبار الفرضية بأن $(b = 0.75)$. يفترض أن نقوم بتقدير b ، وبعدئذ ، نرى إلى أي مدى يكون هذا التقدير قريبا من 0.75 . فإذا كان الاختلاف بين تقديرنا وبين 0.75 «صغيرا» جدا فإننا سنشعر بأن النتائج تدعم الفرضية . ولكن من الناحية الأخرى إذا كان تقديرنا يختلف عن 0.75 بدرجة «كبيرة» فإننا سوف نستنتج أن الفرضية لم تتأكد بنتائجنا المشاهدة ، وسيكون لدينا سبب جيد للاعتقاد بأن $b \neq 0.75$. وحتى يمكننا بناء تحليل ذي معنى من هذا النوع ، ينبغي أن نكون قادرين على التمييز بين الاختلافات «الصغيرة» و «الكبيرة» بين القيم المقترضة والقيم المقدرة للمعلمة . وعلى سبيل فكرة تمهيدية عامة ، فإننا نميز بين ماهو كبير وماهو صغير من هذه الاختلافات استنادا إلى بناء فترة ثقة للمعلمة موضع السؤال وملاحظة هل تكون القيمة المقترضة للمعلمة تقع داخل هذه الفترة أم لا . وعلى سبيل المثال ، إذا كان الاحتمال هو 0.95 أن الفترة $(\hat{b} + 0.07)$ تحتوي على b ، فإننا سوف نرفض الفرضية بدرجة ثقة 95% بأن $(b = 0.75)$ إذا كانت b ، بناء على بياناتنا ، قدرت بـ 0.89 . وطالما أن اتساع الفترة يرتبط بالثقة التي نشعر بها حول احتوائها على قيمة المعلمة ، فإنه يتبع ذلك أن درجة الثقة التي نشعر بها في نتائج اختبارنا للفرضيات يعتمد مباشرة على الاحتمال المرتبط بفترتنا للثقة .

افتراض إضافي

لكي يمكننا اختبار الفرضيات وبناء فترات الثقة لقيم المعلومات في نموذج الانحدار ، ينبغي علينا أولا أن نضيف المزيد لـ خصائص الخطأ العشوائي u_i في نموذجنا . ففي الفصل الثاني وضعنا أربعة فروض حول خصائص الأخطاء العشوائية : فهي متغيرات عشوائية ذات قيم متوقعة صفرية ، ولها التباين نفسه ، ولها تغايرات صفرية وأخيرا هي مستقلة عن المتغيرات الموجودة في الطرف الأيمن من المعادلة . والآن ، فمن المفيد أن نضيف افتراضا خامسا : سنفترض أن الأخطاء العشوائية

موزعة توزيعاً طبيعياً، أي أن كثافتها، أو دالة احتمالها، هو المنحنى الطبيعي. ولما كان التوزيع الطبيعي له معلمتان فقط الوسط الحسابي والتباين، فإن المتغير الذي يكون موزعاً توزيعاً طبيعياً يمكن تحديده بوسطة الحسابي وتباينه. ويمكننا أن نلخص بعض الافتراضات المرتبطة بالخطأ العشوائي u_i بوسطة الرموز $N(0, \sigma^2)$ وبالكلمات، تشير $N(0, \sigma_u^2)$ إلى متغير موزع توزيعاً طبيعياً بوسط صفر وتباين σ_u^2 .*

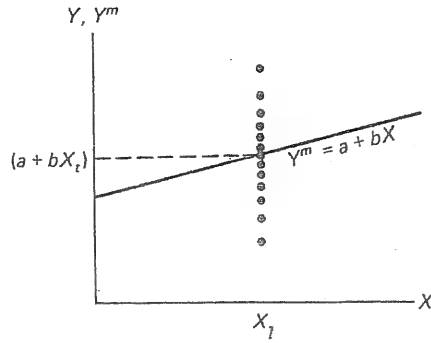
وباستخدام هذا الافتراض الإضافي الخاص بطبيعة الخطأ العشوائي يمكننا أن نوضح بالرسم السمات الأساسية لنموذج الانحدار كما في الشكل (٣-١). فلأي قيمة معطاة من X ، X_i ، مثلاً تكون القيمة المتوسطة لـ Y هي $(a+bX_i)$ ولكن لوجود الخطأ العشوائي فإنه لا يمكن تحديد Y تماماً بوسطة وسطه الحسابي. فإذا كان لدينا مشاهدات متكررة عن Y تناظر جميعها القيمة المعينة لـ X_i فلن نتوقع تساوي قيم Y المشاهدة بقيمتها المتوسطة $(a+bX_i)$. فضلاً عن ذلك، وطالما أن السبب الوحيد لانحرافات القيم المشاهدة عن وسطها الحسابي هو الخطأ العشوائي، فإنه يترتب على ذلك أن تتوزع هذه الانحرافات توزيعاً طبيعياً إذا كان الخطأ العشوائي موزعاً توزيعاً طبيعياً كذلك. على سبيل المثال، إذا كان لدينا شكل لانتشار المشاهدات المكررة لـ Y والمناظرة لـ X_i ، فإننا نتوقع أنها ستشبه انتشار النقاط في الشكل (٣-١) حيث تكون البيانات (المشاهدات) أكثر كثافة بالقرب من القيمة المتوسطة لـ Y ، $(a+bX_i)$ ، عنه بعيداً عنها. وسبب ذلك هو أن ارتفاع منحنى الكثافة الطبيعي يتناقص تدريجياً كلما تحركنا بعيداً عن قيمته المتوسطة (حيث تكون القيمة المتوسطة في هذه الحال هي الصفر).**

* $N(\mu_x, \sigma_x^2)$ هو الرمز المعتاد الذي يشير إلى أن المتغير X موزع توزيعاً طبيعياً بوسطه μ_x وتباينه σ_x^2 .

وينبغي عدم الخلط بين N هذه مع n الصغيرة التي تشير إلى حجم العينة.

** لتبسيط العرض، سنفترض، من الآن فصاعداً في هذا الكتاب، أن الخطأ العشوائي موزع توزيعاً طبيعياً.

ولكن كثيراً من النتائج المذكورة في هذا الكتاب لا تتطلب من الناحية الفنية هذا الافتراض.



شكل (١-٣)

دعنا الآن نتجه إلى مقدراتنا، \hat{a} و \hat{b} ، لمعلمات نموذج الانحدار. تذكر أنه عندما قمنا بإيجاد تباينات هذه المقدرات من الفصل الثاني أوضحنا أن \hat{a} و \hat{b} توليفات خطية من الأخطاء العشوائية u_1, u_2, \dots, u_n وذلك لقيم معطاة من المتغير المستقل. وتوجد نظرية في الإحصاء تنص على أن التوليفات الخطية من متغيرات طبيعية تكون، أيضاً، موزعة توزيعاً طبيعياً. وينتج عن ذلك أنه لأي مجموعة من قيم X_i ، فإن كلا من \hat{a} و \hat{b} ينبغي أن تكون موزعة توزيعاً طبيعياً. وقد وجدنا في الفصل الثاني أن القيم المتوسطة لكل من \hat{a} و \hat{b} هي a و b على الترتيب. كما أوجدنا صيغاً (في 2.87 و 2.96) لتبايناتها. وباستخدام هذه النتائج مضافاً إليها افتراضنا الإضافي بأن u موزعة توزيعاً طبيعياً، فإنه يمكننا أن نستنتج أن:

$$\hat{a} \text{ is } N\left(a, \frac{\sigma_u^2 \sum X_i^2}{n \sum (X_i - \bar{X})^2}\right) \quad (3.1)$$

$$\hat{b} \text{ is } N\left(b, \frac{\sigma_u^2}{\sum (X_i - \bar{X})^2}\right) \quad (3.2)$$

هذه النتيجة مفيدة خاصة، لأنه إذا كانت مقدراتنا \hat{a} و \hat{b} موزعة توزيعاً طبيعياً فإنه يمكننا استخدام طرق الإحصاء المعتادة لاختبار الفرضيات المرتبطة بـ a و b . وكما نتذكر من مبادئ الإحصاء الأولية، أنه إذا كان لدينا متغير (مثلاً v) موزع توزيعاً طبيعياً بوسط حسابي μ_v وتباين σ_v^2 [أي إذا كان $v \sim N(\mu_v, \sigma_v^2)$] فإن:

$$Z = \frac{v - \mu_v}{\sigma_v} \quad (3.3)$$

سيكون متغيراً طبيعياً معيارياً، وبمعنى آخر، فإن $Z \sim N(0, 1)$. ويمكننا، باستخدام جدول التوزيع الطبيعي، المعيارى أن نتذكر بعض النتائج الاحتمالية حول قيمة Z . في مثل هذا الجدول، نجد، مثلاً أن*

$$\begin{aligned} \text{Prob}(-1.65 \leq Z \leq 1.65) &= 0.90 \\ \text{Prob}(-1.96 \leq Z \leq 1.96) &= 0.95 \\ \text{Prob}(-2.58 \leq Z \leq 2.58) &= 0.99 \end{aligned} \quad (3.4)$$

وهذا المثال يعني أنه إذا اختيرت قيمة Z عشوائياً، فإن احتمال أن تكون قيمة Z محصورة بين (-1.65) و $(+1.655)$ سيكون 90. لاحظ أن العبارة الأولى في (3.4) تتضمن كلا من:

$$\text{Prob}(Z \leq -1.65) = 0.05 \quad (3.5)$$

و

$$\text{Prob}(Z \geq 1.65) = 0.05 \quad (3.6)$$

ومرة أخرى، تذكر أن سبب ذلك هو تماثل المنحنى الطبيعي، لذلك، فإذا كان 0.90 من المساحة يقع بين ± 1.65 فإن 0.05 من تلك المساحة ينبغي أن تقع إلى يسار -1.65 و 0.05 إلى يمين +1.65.

افترض الآن أن σ_v^2 معلومة. فإذا رمزنا إلى تباينات \hat{a} و \hat{b} المعطاة في

* انظر الجدول الإحصائي الأول في نهاية الكتاب.

المعادلتين (3.1) و (3.2) بالرموز σ_a^2 و σ_b^2 على الترتيب يصبح لدينا :

$$\left(\frac{\hat{a} - a}{\sigma_{\hat{a}}} \right) \sim N(0,1) \text{ و } \left(\frac{\hat{b} - b}{\sigma_{\hat{b}}} \right) \sim N(0,1) \quad (3.7)$$

حيث إن \hat{a} و \hat{b} هي الانحرافات المعيارية لكل من a و b على الترتيب. ويمكننا في ضوء المعادلة (3.7) أن نكون النتائج نفسها حول $(\hat{a} - a)/\sigma_{\hat{a}}$ و $(\hat{b} - b)/\sigma_{\hat{b}}$ التي كونها حول المتغير الطبيعي المعياري Z في المعادلة (3.3)، على سبيل المثال، فإن :

$$\text{Prob} \left(-1.96 \leq \frac{\hat{b} - b}{\sigma_{\hat{b}}} \leq 1.96 \right) = 0.95 \quad (3.8)$$

اختبار $b = b_0$ مقابل $b \neq b_0$ ، مع معرفة σ_b^2

في مقدورنا الآن-أن نبني فترات ثقة وأن نستخدمها لاختبار فرضياتنا. فبالنسبة لـ \hat{b} ، مثلاً، يمكننا إعادة ترتيب الحدود الموجودة في المعادلة (3.8) للحصول على :

$$\text{Prob}(\hat{b} - 1.96\sigma_{\hat{b}} \leq b \leq \hat{b} + 1.96\sigma_{\hat{b}}) = 0.95 \quad (3.9)$$

وتبين المعادلة (3.9) أنه، باحتمال قدره 0.95، تشتمل الفترة :

$$(\hat{b} - 1.96\sigma_{\hat{b}}, \hat{b} + 1.96\sigma_{\hat{b}}) \quad (3.10)$$

على قيمة b ، لذلك تكون المعادلة (3.10) هي فترة ثقة 95% لـ b . ويصبح منهج الاختبار المقترح على النحو التالي : نحسب، على أساس عيبتنا، كلا من \hat{b} ، $\sigma_{\hat{b}}$. بعد ذلك نحسب كلا من $(\hat{b} - 1.96\sigma_{\hat{b}})$ و $(\hat{b} + 1.96\sigma_{\hat{b}})$. وأخيراً نتبين هل القيمة المفترضة لـ b تقع داخل الفترة المعطاة في المعادلة (3.10) أم لا. فإذا لم تكن كذلك نرفض فرضيتنا بدرجة ثقة 95%. أما إذا كانت تقع في تلك الفترة فإننا نصرح بأن البيانات تتسق مع فرضيتنا ولذا نقبلها.

ينبغي علينا أن ندرك أن طريقة اختبارنا ليست معصومة من الخطأ، فمثلاً، تتضمن المعادلة (3.9) وجود فرصة تعادل 5% بأن الفترة $(\hat{b} \pm 1.96 \sigma_{\hat{b}})$ لا تحتوي على b . وهكذا، فقد نرفض الفرضية المرتبطة بقيمة b (مثلاً $b = 0.75$) حتى ولو كانت فرضية صحيحة. وبدقة أكثر قد نرفض فرضيتنا المسبقة (فرضية العدم) عندما تكون هذه الفرضية صحيحة في الحقيقة. يطلق على هذا الشكل من الخطأ خطأ من النوع الأول (Type I error)، ويرمز لاحتمال الوقوع في هذا الخطأ، عادة، بالرمز α ، ويطلق عليه مستوى المعنوية. وفي مثالنا، نجد $(\alpha = 0.05)$. وبهذه المناسبة، يشار إلى α ، أيضاً، بـ «حجم» الخطأ من النوع الأول.

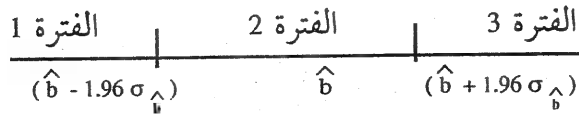
افترض مرة أخرى أن فرضيتنا للعدم (H_0) هي $b = 0.75$. فإذا رفضنا $H_0 : b = 0.75$ وإذا لم تتوافر لنا معلومات إضافية، فمن الواضح أنه لم يبق لنا سوى القول بأن $b \neq 0.75$ ويشار إلى هذه العبارة في الإحصاء بالفرضية البدئية للفرضية H_0 ويشار إليه عادة بالرمز H_1 . وبمعنى آخر تصبح العبارة الكاملة لفرضيتنا محل الاهتمام هي $H_0 : b = 0.75$ و $H_1 : b \neq 0.75$ أي أن منهجنا في الاختبار سيؤدي إما إلى H_0 أو H_1 .

والخطأ من النوع الأول ليس هو النوع الوحيد من الأخطاء الذي يمكن أن نقع فيه، فمثلاً، قد نقبل H_0 حتى إذا كان غير صحيح (أي حتى إذا كان H_1 هي الفرضية الصحيحة). وعلى سبيل المثال، افترض أن فترتنا للثقة أصبحت من (0.86 إلى 0.92) افترض أيضاً أن $H_0 : b = 0.87$ ولكن القيمة الحقيقية لـ b هي 0.88، حينئذ، فإنه، طالما أن 0.87 تقع في داخل فترتنا فقد نقبل $H_0 : b = 0.87$. ولكن في هذه الحال فإن $H_1 : b \neq 0.87$ هو الصحيح، ومثل هذا الاحتمال يوجد، دائماً، طالما أن فترتنا للثقة تشتمل على أكثر من قيمة واحدة. ويطلق على هذا النوع من الخطأ (أي قبول فرضية العدم عندما تكون خطأ) الخطأ من النوع الثاني (Type II error) ويشار، عادة، إلى احتمال الوقوع في النوع الثاني من الخطأ (والذي يشار إليه أيضاً بـ «بحجم» النوع الثاني من الخطأ) بالرمز β .

ويمكننا أن ندرك أن احتمال الوقوع في النوع الثاني من الخطأ ينخفض عندما تكون القيمة الحقيقية للمعلمة تختلف اختلافا كبيرا عن قيمتها المفترضة. فمثلا، إذا كان $H_0: b=0.75$ ولكن في الحقيقة $b=0.98$ فإن تقديرنا لـ b (\hat{b}) سيكون قريبا من 0.98، لذلك تحتوي فترتنا للثقة ($\hat{b} \pm 1.96 \sigma_{\hat{b}}$) على مدى من القيم مركزة (تركيزا تقريبا) حول (0.98). فإذا كان ذلك صحيحا فسوف ننتهي برفض فرضيتنا بأن، $b=0.75$. ومن الناحية الأخرى، إذا كانت قيمة $b=0.751$ فإن فترة الثقة ستشمل عادة، على قيمتنا المفترضة 0.75. وهكذا، فستكون هناك فرصة كبيرة أن نقع في الخطأ من النوع الثاني. وهكذا يتضح لنا أننا سنقع على الأرجح في النوع الثاني من الخطأ عندما تكون القيمة المفترضة للمعلمة «قريبة» من القيمة الحقيقية. باختصار، يصعب وضع حدود فاصلة دقيقة بين الفروض المرتبطة بمعلمات النموذج.

اختبار الفرضيات : تفسير

الخلاصة أن منهجنا للاختبار يتلخص في قبول فرضية العدم أو رفضها على أساس الفرق بين القيمة المفترضة للمعلمة وتقديرنا لها. وحتى يتضح ذلك لنا أكثر بتفصيل اعتبر فترة الثقة 95% لـ b مثل الفترة 2 في الشكل (٣-٢). فإذا كان الفرق بين \hat{b} والقيمة المفترضة لـ b ، مثلا b_0 (مثل $b_0 = 0.75$) كبير جدا بحيث يزيد على $1.96 \sigma_{\hat{b}}$ حيثئذ، فإن b_0 ستقع إما في الفترة 1 أو الفترة 3. وفي مثل هذه الحالة سنرفض فرضية لعدم بأن $b = b_0$ عند درجة ثقة 95% وهكذا نجد أن ما يحدد مانطلق عليه فرقا Discrepancy «كبيراً» مقارنة بفرق «صغير» هو، ببساطة، مضاعف الانحراف المعياري المقدّر ($\text{أي } 1.96 \sigma_{\hat{b}}$). ويبدو هذا منطقيا. فإذا كانت $\sigma_{\hat{b}}$ مثلا، كبيرة فإن دقة مقدارنا ستكون منخفضة، ومن ثم، سنقبل بوجود فرق كبير بين تقديرنا والقيمة المفترضة للمعلمة قبل رفض فرضية العدم.



شكل (٣-٢)

مناطق القبول والرفض

رأينا في المثال السابق أن فرضية العدم ستقبل إذا كانت :

$$|\hat{b} - b_0| < 1.96\sigma_{\hat{b}} \quad (3.11)$$

وبإعادة كتابة المعادلة (3.11) في صورة أخرى، يتبين لنا أن قبول فرضية العدم يرتبط بالمدى التالي لقيم \hat{b} :

$$b_0 - 1.96\sigma_{\hat{b}} < \hat{b} < b_0 + 1.96\sigma_{\hat{b}} \quad (3.12)$$

ويطلق على مدى القيم لـ \hat{b} والمحددة في المعادلة (3.12) منطقة القبول acceptance region ومنطقة القبول (عموما) هي ذلك المدى من القيم لمقدرنا الذي يؤدي إلى قبول فرضية العدم. وعلى العكس، يطلق على مدى القيم لمقدرنا الذي يؤدي إلى رفض فرضية العدم منطقة الرفض rejection region أو في بعض الأحيان المنطقة الحرجة critical region. وفي المثال السابق تكون المنطقة الحرجة هي :

$$\begin{aligned} \hat{b} &> b_0 + 1.96\sigma_{\hat{b}} \\ \hat{b} &< b_0 - 1.96\sigma_{\hat{b}} \end{aligned} \quad (3.13)$$

ومن الواضح الآن، أنه يمكن اختبار الفرضيات بواسطة منهج مماثل يتكون أولا من تحديد مناطق القبول والرفض، وبعد ذلك (وعلى أساس العينة)، نحدد أي المناطق يحتوي على تقدير معلمتنا.

فترات الثقة : تفسير

قبل أن نمضي في التحليل، علينا مناقشة المعادلة (3.9) بشئ من التفصيل. فعلى افتراض أن قيم المتغير المستقل معطاة، فإن المتغير العشوائي في العبارة الاحتمالية هو \hat{b} . لاحظ أن b و $\sigma_{\hat{b}}$ ثابت، أي أنهما ليسا متغيرين عشوائيين. وتقول المعادلة (3.9) إن احتمال أن تحتوي الفترة $(\hat{b} - 1.96 \sigma_{\hat{b}}, \hat{b} + 1.96 \sigma_{\hat{b}})$ على b هو 0.95، وبمعنى آخر، افترض أنه يمكننا الحصول على عدد كبير من العينات (1000 عينة، مثلاً) كل منهم بحجم 50. افترض، أيضاً إن قيمة المتغير المستقل X_i 's لم تتغير في جميع هذه العينات. وافترض، أخيراً، أننا حسبنا \hat{b} لكل عينة. وطالما أن الأخطاء العشوائية يتوقع أن تختلف من عينة لأخرى فإن قيم Y_i 's سوف تختلف بين العينات حتى ولو لم تتغير قيم X_i 's وطالما أن \hat{b} تعتمد على Y_i 's فإنها ستختلف من عينة لأخرى. فإذا حسبنا الفترة $(\hat{b} \pm 1.96 \sigma_{\hat{b}})$ لكل عينة، فإنه سيكون لدينا 1000 فترة مختلفة. وفي الحقيقة، فإن جوهر المعادلة (3.9) هو أننا، نتوقع بهذه الطريقة أن $0.95(1000)=950$ من هذه الفترات سيحتوي على الثابت b .

بعض الملاحظات على الأخطاء من النوع الأول والنوع الثاني

بنينا اختباراتنا السابقة على أساس فترة ثقة 95% ومن ثم، مستوى معنوية 5%، ولكن الباحث، في الحقيقة، هو الذي يختار مستوى المعنوية، ولذا، يمكننا، بسهولة، (متى كان ذلك مرغوباً فيه) أن نبني اختباراتنا باستعمال مستويات أخرى من المعنوية. فإذا أردنا، مثلاً، أن نتأكد بدرجة عالية أننا لن نرفض الفرضية $b=0.75$ عندما تكون في الحقيقة صحيحة، علينا أن نبني اختباراً ذا احتمال أصغر لوقوع الخطأ من النوع الأول $\alpha < 0.05$ ونختار α في مثل هذه الحالة معادلة لـ 0.01، على الرغم من أن الباحث حر في اختيار أي قيمة لـ α . افترض لاكمال مثالنا السابق، أننا نرغب في بناء اختبار حيث يكون $\alpha < 0.01$ حيثئذ سوف نختار 0.99 مستوى احتمال في المعادلة (3.8) وسوف ننتهي بالفترة.

$$(\hat{b} \pm 2.58\sigma_{\hat{b}}) \quad (3.14)$$

ومرة أخرى، سوف نقبل أو نرفض فرضيتنا معتمدين على ما إذا كانت الفترة قد اشتملت على القيمة المفترضة لـ b أو لم تشتمل عليها.

وتصبح فترة الثقة بالمعادلة (3.14) المرتبطة بخطأ من النوع الأول حجمه $\alpha = 0.01$ أكبر من الفترة بالمعادلة (3.10) التي يكون حجم خطأها من النوع الأول $\alpha = 0.05$. ومن الواضح أنه إذا كانت فترتنا للثقة أكبر ولا تزال مركزة عند مقدارنا \hat{b} نفسه، فإن احتمال أن تقع في الخطأ من النوع الثاني يكون كبيراً. أي أنه حتى إذا كان H_0 غير صحيح، فإننا، على الأرجح، سنقبله إذا كانت فترة الثقة كبيرة عما لو كانت صغيرة. ويشير هذا إلى أنه، عندما نقوم بتخفيض احتمال الوقوع في الخطأ من النوع الأول فإننا نزيد في الوقت نفسه احتمال الوقوع في الخطأ من النوع الثاني. هذا التبادل trade off معروف جيداً في الإحصاء، لأنه، عموماً، (مع وجود عينة معطاة) لا يمكن أن نقلل من احتمال الوقوع في الخطأين الأول والثاني في الوقت نفسه. وباختصار، فإن أحدهما يمكن أن يقلل على حساب الآخر. وعلى ضوء هذه المناقشة نجد أنه، كي نبني فترة ثقة لأغراض اختبار الفرضيات، فإن علينا أولاً أن نختار احتمال ارتكاب (الوقوع) في الخطأ من النوع الأول أو الخطأ من النوع الثاني. وفي العادة، يختبر الاقتصاديون الفرضيات المرتبطة بقيم معينة للمعلمة ($b=0.75$) عن طريق اختبار α عند 0.05 أو 0.01، وبالطبع، فإنهم لا يوجهون اهتمامهم نحو حجم الخطأ من النوع الثاني المرتبط بذلك الاختبار.

وليس هذا بالطبع هو أفضل الحلول، فإن حجم الخطأ من النوع الأول الذي يحدد، بدوره، حجم الخطأ من النوع الثاني ينبغي اختياره بعناية. والنقطة المهمة التي يجب أن نتذكرها دائماً هي أن أي من الخطأين قد يؤدي إلى القيام بعمل (أو اتخاذ قرار) له نتائج غير مرغوب فيها. لذلك، فقد نفكر في الخسارة المرتبطة بكل نوع من الأخطاء ونختار تلك الخسارة الأقل. ينبغي علينا أن نوزن، بطريق أو

بآخر، أهمية هذه الخسائر في تحديد حجم الخطأ من النوع الأول. افترض (على سبيل المثال) أن الحكومة تقوم بعمل دراسة لأسباب الشغب. افترض أن فرضية العدم هي أن سياسة حكومية معينة لا تأثير لها في تخفيض المؤثرات (أو الأسباب) التي تؤدي إلى الشغب. افترض أن الفرضية البديلة هي أنها تؤثر. في هذا المثال يؤدي الخطأ من النوع الأول إلى توقع أن السياسة الحكومية تكون مؤثرة في الوقت الذي لا تكون فيه في الواقع كذلك. وأحد النتائج المترتبة على الوقوع في هذا الخطأ هو أن الاعتمادات المالية قد تهدر على سياسة حكومية غير فعالة. ومن الناحية الأخرى، يؤدي الوقوع في الخطأ من النوع الثاني إلى الاستنتاج بأن السياسات الحكومية لا تعمل في الوقت الذي تقلل فيه، فعلا من احتمالات وقوع الشغب. ونتيجة هذا الخطأ هو أن الاعتمادات لن تنفق على سياسة فعالة. ومن الواضح أن تقويم مدى أهمية كل من هذه الأخطاء سيعتمد على بعض الافتراضات الإضافية. افترض على سبيل المثال، أننا نعتبر عددا من تلك السياسات لمواجهة الشغب، إلا أنه بسبب الاعتمادات المحدودة، لا يمكن تنفيذ سوى سياسة واحدة. في هذه الحال من الواضح أن الخطأ من النوع الأول سيكون مهما جدا لأنه سيؤدي إلى هدر الاعتمادات المحدودة في بيئة محفزة للشغب. بينما يتطلب الخطأ من النوع الثاني، ببساطة، أن يعتبر الباحثون سياسة أخرى.* وهكذا ففي هذا المثال، قد نضع α عند مستوى منخفض جدا (ربما أقل من 0.01).

وعلى النقيض من ذلك، افترض أن الاعتمادات متوافرة نسبيا وأن واحدة، فقط، من تلك السياسات تعد سياسة فعالة. افترض، أيضا، أن الحكومة (وبسبب توافر الاعتمادات) مستعدة لتنفيذ السياسات كافة التي تبدو واعدة. في هذه الحالة فإن تبديد الاعتمادات على سياسة غير فعالة بسبب الخطأ من النوع الأول قد لا يكون خطيرا جدا حيث تصبح الخسارة الوحيدة هي تبديد الاعتمادات ذاتها. ونتيجة لذلك، فإن الحكومة قد ترفض سياسة فعالة ومن ثم، قد تحدث الاضطرابات

* نفترض هنا أن السياسات الأخرى الكفاء موجودة بالفعل وسيتم اعتبارها في فترة زمنية «معقولة» من الزمن.

الاجتماعية. في مثل هذه الحال يكون الخطأ من النوع الثاني أكثر أهمية من الخطأ من النوع الأول. لذلك، فإنه في مثل هذه الحالات، سيكون التحديد الصحيح لحجم الخطأ من النوع الأول أكثر ارتفاعاً عن $(\alpha=0.05)$ ، فمثلاً، قد نجعل $\alpha=0.30$ أو أكبر من ذلك من أجل تقليل حجم الخطأ من النوع الثاني الأكثر تكلفة. ونود بهذه المناسبة أن نشير إلى أن الحل الدقيق لمشكلة تحديد α معقد جداً وهو، في الحقيقة، يتعدى جداً عن مجال هذا الكتاب.* ولكننا نأمل أن يكون لديك الآن تفهم أفضل لبعض هذه القضايا في الأقل.

الفرضية: $b \neq 0$

عند التطبيق، لا يكون لدى الاقتصاديين، عادة، افتراضات تحدد قيماً معينة للمعاملات. وتقتصر النظرية الاقتصادية، غالباً، وجود علاقة موجبة أو سالبة بين متغيرين. وهنا نجد أن الافتراضات التي تهتم الاقتصاديين، غالباً، ماتكون، فقط، تحديد إشارة معلمة معينة. وفي الحقيقة، تشير النظرية الاقتصادية في بعض الحالات، فقط، إلى ماهية المتغيرات المرتبطة ببعضها بعضاً، ولكن، هل هذه العلاقة موجبة أو سالبة، تعد مسألة تحتاج إلى الفحص التطبيقي. افترض، على سبيل المثال، أننا نهتم بكيفية تغير استهلاك البطاطس عند تغير مستوى دخل الأسرة. وفرضيتنا هي أن :

$$Q_t = a + bY_t + u_t \quad (3.15)$$

حيث Q_t هي كمية البطاطس المستهلكة بواسطة الأسرة رقم t ، Y_t هو مستوى دخل الأسر و u_t هو الخطأ العشوائي. من غير الواضح في هذه الحالة توقع إشارة b (هل تكون موجبة أم سالبة). وبالنسبة لغالبية السلع، نتوقع أن تتزايد الكمية المستهلكة مع الدخل. ولكن هناك فئة من السلع يطلق عليها الاقتصاديون «الدنيا» وهي التي

* لمعرفة المزيد عن هذا الموضوع يرجع إلى:

Chapter 12 of Alexander H. Mood and Franklin and A. Graybill. *An Introduction to the Theory of Statistics* (New York: McGraw-Hill, 1963); and Chapters 1 and 2 in Arnold Zellner. *An Introduction to Bayesian Inferences in Econometrics* (New York: Wiley, 1971).

يتغير استهلاكها عكسيا مع مستوى الدخل . والفكرة المتضمنة هنا هي أن زيادة دخل الأسرة يمكنها من شراء قائمة أكثر تنوعا وأكثر تكلفة من الطعام . ومن غير المحتمل أن نندهش كثيرا إذا وجدنا الأسر مرتفعة الدخل تحل بعض السلع الأخرى من الطعام محل البطاطس . ومن ناحية أخرى ، يمكن أن تستهلك الأسر الغنية ، في الحقيقة ، كميات من البطاطس أكثر من الأسر الأقل دخلا . وفي هذه الحال ، نجد أنه على الرغم من أننا نعتقد بوجود علاقة ما بين Q و Y ، فإننا لسنا متأكدين من طبيعة (إشارة) هذه العلاقة . وهكذا تصبح الفرضية التي نريد اختبارها في هذه الحال ، ببساطة ، $b \neq 0$ بدون أية قيود مسبقة على إشارة b .

كيف نختبر الفرضية $b \neq 0$ ؟ من الواضح أنه لا يمكننا أن نبني ببساطة (على سبيل المثال) فترة ثقة 95% ثم ، نفحص بعد ذلك هل تتضمن هذه الفترة أي قيمة من قيم $b = 0$. فإذا فعلنا ذلك فإننا سنقبل دائما فرضيتنا ، وذلك بسبب أن فترتنا ستضمن دائما بعض القيم التي يكون فيها $b \neq 0$. ويصبح احتمال الوقوع في الخطأ من النوع الثاني في هذه الحال ، هو الواحد الصحيح .

ومن الواضح ، أنه لا يمكن أن يكون لدينا أي ثقة في الفرضيات التي نقبلها إذا كان منهج الاختبار يتضمن دائما عدم رفضها . ومن أجل تصحيح ذلك ، يختبر الاقتصاديون الفرضيات من النوع $b = 0$ عن طريق وضع حجم الخطأ من النوع الثاني مساويا لصغير الحجم ، عادة ، 0.05 أو 0.01 ويتم ذلك ، بسهولة ، عن طريق إعادة صياغة relabeling . رقما الفرضيات . على سبيل المثال ، فإن الفرضية البديلة لفرضية العدم $b \neq 0$ هي $b = 0$. لذلك ، يكون الخطأ من النوع الثاني هو قبول $b \neq 0$ عندما يكون في الحقيقة $b = 0$. افترض الآن أننا نعتبر $b = 0$ فرضيتنا للعدم و $b \neq 0$ هي الفرضية البديلة . وأنا اخترنا مستوى المعنوية $\alpha = 0.05$. يدل اختبارنا أن احتمال رفض $b = 0$ عندما يكون في الحقيقة $b = 0$ هو $\alpha = 0.05$. ولما كان رفض $b = 0$ يعادل قبول $b \neq 0$ فإن نتيجتنا تتحقق (وكما هو مطلوب) ، أي أننا بنينا اختبارنا بحيث يكون احتمال قبول الفرضية $b = 0$ عندما تكون في الحقيقة $b = 0$ هو $\alpha = 0.05$. على سبيل المثال ، إذا لم تكن هناك علاقة بين كمية البطاطس المستهلكة Q_1

ومستوى دخل الأسرة Y_t ، فإن الاحتمال سيكون 0.05. أي أن منهجنا في الاختبار سيؤدي بنا إلى الاعتقاد بوجود علاقة $b = 0$ بين هذه المتغيرات. ولتلخيص ماسبق لاختبار الفرضية بأن قيمة معلمة معينة (مثلا b) ليست صفرا نبنى الفرضيات.

$$H_0: b = 0, \quad H_1: b \neq 0, \quad (3.16)$$

ونختار α طبقا لمستوى الثقة المرغوب فيه. وكما ذكرنا، يحدد الاقتصاديون، عادة، α عند 0.05 أو 0.01. فإذا رفضنا $H_0: b = 0$ نصرح بأن تقديرنا مختلف معنويا عن الصفر، أما إذا قبلنا $H_0: b = 0$ فإننا نذكر أن تقديرنا لا يختلف معنويا عن الصفر. لاحظ أن الحالة الأخيرة تعني أننا، في الحقيقة، غير قادرين على إيجاد علاقة منتظمة بين المتغيرات تقابل مستويات المعنوية تم اختيرت.

الفرضيات $b < 0$ و $b > 0$

اختبرنا (في المثال السابق) الفرضيات $H_0: b = b_0$ و $H_1: b \neq b_0$ حيث إن b_0 قد تكون صفرا. يطلق على مثل هذا الاختبار الاختبار ذو الذيلين two tailed test. أي أن الفرضية البديلة H_1 هي b إما أن تكون أكبر من b_0 أو أقل منها. في ظل منهجنا للاختبار تؤدي الانحرافات الكبيرة الموجبة أو السالبة لـ \hat{b} عن القيمة المحددة لـ b_0 إلى رفض فرضية العدم. وبالمقابل، يهتم الاقتصاديون كثيرا باختبار الدليل الواحد. وطالما أن النظرية الاقتصادية تقترح إشارة معينة للعلاقة بين المتغيرات، فإن الفرضيات المشتقة من النظرية تكون في الشكل $b > 0$ أو $b < 0$. افترض (مرة أخرى) أننا نرغب في بناء فترة ثقة 95% للفرضيات التي نقبلها. ستكون، حيثئذ، فرضياتنا على النحو $H_0: b = 0$ مقابل $H_1: b > 0$ أو $H_1: b < 0$ بحيث يكون احتمال الوقوع في الخطأ من النوع الأول هو 0.05.

وتنفيذ الاختبارات ذات الدليل الواحد واضح ومباشر. اعتبر، على سبيل

المثال، الفرضيات:

$$H_0: b = 0, \quad H_1: b > 0, \quad (3.17)$$

تنص هذه الفرضيات على أن b إما مساوية للصفر أو أكبر منه، وهكذا لا ينبغي أن

نشغل أنفسنا بالقيم السالبة لـ b . وعلى افتراض أننا نرغب في بناء فترة ثقة 95% لفرضيتنا المقبولة، يكون منهجنا في اختبار الفرضيات هو بناء حد أدنى Lower bound لقيمة b باحتمال 95% أن يكون ذلك الحد الأدنى أقل من b . هذا الحد الأدنى يزودنا فعلا بفترة ثقة 95% مفتوحة Open ended. وكما عرفنا من قبل تعتمد قيمة هذا الحد الأدنى على قيمة مقدرها \hat{b} . وسيكون منهجنا للاختبار هو تحديد قيمة \hat{b} (ومن ثم، حدنا الأدنى) من بيانات العينة ثم، تحديد هل يكون ذلك الحد الأدنى أكبر من الصفر أم لا ؟ (أي هل يقع الصفر خارج فترتنا للثقة) فإذا كان حدنا الأدنى موجبا فسوف نرفض الفرضية $b=0$ وإذا لم يكن كذلك فسوف نقبل الفرضية ونرفض الفرضية البديلة $b>0$. نشق حدنا الأدنى من العبارة :

$$\text{Prob}\left(\frac{\hat{b}-b}{\sigma_{\hat{b}}} < 1.65\right) = 0.95, \quad (3.18)$$

حيث نجد 1.65 في جدول قيم المنحنى الطبيعي المعياري (انظر الجدول الإحصائي (١) ويمكن إعادة كتابة المعادلة (3.18) على النحو :

$$\text{Prob}(\hat{b} - 1.65\sigma_{\hat{b}} < b) = 0.95 \quad (3.19)$$

وطالما أن \hat{b} (وليست b أو $\sigma_{\hat{b}}$) هي المتغير، نجد من المعادلة (3.19) أن احتمال أن يكون الحد الأدنى $(\hat{b} - 1.65\sigma_{\hat{b}})$ أصغر من قيمة المعلمة b مساويا 0.95. وبناء على ذلك، نرفض فرضية العدم $b=0$ ، وعلى أساس معلومات العينة، إذا كانت $\hat{b} - 1.65\sigma_{\hat{b}}$ أكبر من الصفر. وبالمقابل، إذا كانت $(\hat{b} - 1.65\sigma_{\hat{b}}) \leq 0$ فسنقبل الفرضية $b=0$. وفي حالتنا الأخيرة هذه نقول إن تديرنا ليس مختلفا معنويا عن الصفر.

في الإحصاء، يطلق على الفترة ذات الشكل $(\hat{b} - 1.65\sigma_{\hat{b}}) < b$ فترة الثقة ذات الطرف الواحد. والآن ينبغي أن تكون قادرا على إثبات أنه، إذا قمنا باختبار

الفرضية $H_0: b=0$ مقابل الفرضية $b < 0$ عند 5% مستوى معنوية، فسوف تنتهي بفترة ثقة 95% ذات ذيل واحد*.

$$b < \hat{b} + 1.65\sigma_{\hat{b}} \quad (3.20)$$

ستشمل في هذه الحالة $(\hat{b} + 1.65\sigma_{\hat{b}})$ حدنا الأعلى لقيمة b ، سنختبر الفرضية $H_0: b=0$ مقابل الفرضية $H_1: b > 0$ عن طريق تقويم \hat{b} من عينتنا ثم، نحدد بعد ذلك هل تكون $(\hat{b} + 1.65\sigma_{\hat{b}})$ أقل من الصفر أم لا. فإذا كانت أقل من الصفر فسوف نرفض الفرضية H_0 أما إذا كانت أكبر من الصفر فسوف نقبلها. ينبغي علينا أن نذكر أخيراً أنه، على الرغم من أننا قد ركزنا المناقشة على المعلمة b فإن مناهج الاختبارات كافة التي ذكرناها لـ b تنطبق، أيضاً، على المعلمة a ، إذ يمكننا استخدام الأساليب نفسها لبناء فترة ثقة ذات ذيل واحد أو ذات ذيلين للمعلمة a لاختبار الفرضيات المرتبطة بقيمة الحد الثابت في معادلة الانحدار.

اختبار الفرضيات، مع عدم معرفة σ_u

افترضنا في تحليلنا السابق أن σ_u^2 معلومة ومن ثم، فإن كلا من $\sigma_{\hat{a}}$ و $\sigma_{\hat{b}}$ معلومان أيضاً، ولكن الحال ليست هكذا، عادة. والآن جاء الوقت الملائم لإسقاط هذا الافتراض. وكما أوضحنا في الفصل الثاني أنه عندما تكون σ_u^2 غير معلومة فإنه ينبغي أن نستخدم مقدرنا لـ $\hat{\sigma}_u^2$ من أجل الحصول على مقدر لتباينات \hat{a} و \hat{b} ، وتكون مقدراتنا للتباين هي:

$$\hat{\sigma}_{\hat{a}}^2 = \frac{\hat{\sigma}_u^2 \sum X_i^2}{n \sum (X_i - \bar{X})^2} \quad \text{و} \quad \hat{\sigma}_{\hat{b}}^2 = \frac{\hat{\sigma}_u^2}{\sum (X_i - \bar{X})^2} \quad (3.21)$$

* إرشاد للحل، نحتاج، في هذه الحالة، إلى حد أعلى. وسنحصل عليه من خلال البدء بتعبير مشابه للموجود في المعادلة (3.18) مع جعل إشارة عدم المساواة معكوسة، وبالتحديد سنبدأ بـ

$$\text{Prob}\left(\frac{\hat{b} - b}{\sigma_{\hat{b}}} > -1.65\right) = 0.95$$

حيث :

$$\hat{\sigma}_u^2 = \frac{\sum (Y_i - \hat{a} - \hat{b}X_i)^2}{(n-2)} = \frac{\sum (Y_i - \hat{Y}_i)^2}{(n-2)}$$

ومع هذا التعديل ، فإنه لم يعد بإمكاننا استخدام المنحنى الطبيعي لاختبار الفرضيات (أو بناء فترة الثقة) المرتبطة بـ a و b . بدلا من ذلك ، يمكننا ، باستخدام المعادلة (3.7) ، أن نكون :

$$\frac{\hat{a} - a}{\hat{\sigma}_a} \text{ و } \frac{\hat{b} - b}{\hat{\sigma}_b} \quad (3.22)$$

حيث إن كلا من التعبيرين في المعادلة (3.22) متغيرات عشوائية يمكن إثبات أن لها توزيع t بـ $(n-2)$ درجات حرية. * ونتبع المنهج نفسه أعلاه باستثناء أننا نستخدم هنا توزيع t بدرجات حرية قدرها $(n-2)$ (انظر الجدول الإحصائي رقم ٢ بدلا من التوزيع الطبيعي لتحديد حدود فترات الثقة).

بعض الأمثلة

لتوضيح منهج الاختبار المبني على توزيع t ، دعنا نعود إلى دالة الاستهلاك التي قدرناها في الفصل الثاني. وباستخدام البيانات عن الإستهلاك والدخل المتاح للسنوات 1960-1969 في الولايات المتحدة الأمريكية وجدنا أن :

$$\hat{a} = 13 \quad \hat{\sigma}_a^2 = 31 \quad (or \hat{\sigma}_a = 5.6)$$

$$\hat{b} = 0.89 \quad \hat{\sigma}_b^2 = 0.0001 \quad (or \hat{\sigma}_b = 0.01)$$

* يشبه توزيع t التوزيع الطبيعي ، فإذا ما تزايدت n وأصبحت كبيرة يؤول توزيع t إلى التوزيع الطبيعي .

دعنا نفترض أننا نرغب في اختبار الفرضية $a > 0$ ، ونرغب أن تكون لدينا 95% فترة ثقة في الفرضية التي نقبلها. * للقيام بذلك، نكون فرضيتنا للعدم على النحو $H_0 : a = 0$ والفرضية البديلة $H_1 : a > 0$ ونجعل مستوى المعنوية عند 0.05. وباستخدام الجدول الإحصائي رقم ٢ لتوزيع t عند 8 (أي 2-10) درجات حرية، نجد أن الحد الأدنى لفترة الثقة ذات الدليل الواحد لـ a هي ** :

$$a > (\hat{a} - t_{n-2; 0.95} \hat{\sigma}_{\hat{a}}) = [13 - 1.8(5.6)] = 2.6$$

طالما أن الحد الأدنى للفترة أعلى من الصف وعند مستوى معنوية 5%، فإننا نقبل $H_1 : a > 0$ أي أننا نستنتج أن قيمة a أكبر من الصفر. لاحظ أنه إذا اخترنا مستوى معنوية 1% فإنه سيكون لدينا الحد الأدنى التالي :

$$a > (\hat{a} - t_{n-2; 0.99} \hat{\sigma}_{\hat{a}}) = [13 - 2.99(5.6)] = -3.2$$

وسيمدنا هذا بفترة الثقة $(a > -3.2)$ التي تشتمل على الصفر. ولذا تم قبول $H_0 : a = 0$ ومن ثم، استنتاج أن a تساوي الصفر. وهذا يشير إلى أنه ينبغي التفكير ملياً في حجم الخطأ من النوع الأول، لأن نتائج اختبارنا تعتمد عليه. ***
دعنا أخيراً نبني فترة ثقة 99% ذات طرفين لـ b . وجدنا من الجدول الإحصائي رقم (٢) أن الفترة هي :

$$(\hat{b} \pm t_{n-2; 0.995} \hat{\sigma}_{\hat{b}}) = 0.89 \pm 3.36(0.01) = (0.89 \pm 0.03)$$

تشتمل هذه الفترة على مدى من القيم للميل الحدي للاستهلاك MPC من -0.92 إلى 0.86، فإذا قمنا باختبار فرضية العدم $H_0 : b = 0.75$ مقابل $H_1 : b \neq 0.75$ عند 1% مستوى معنوية فسرفض H_0 .

* ملاحظة : ينبغي أن نبني الفرضيات قبل تحليل البيانات وقبل الوصول إلى التقديرات، فإذا لم يتحقق ذلك فسنعاني مشكلة المنطق الدائري circular reasoning.

** الترميز المستخدم هنا هو الترميز الشائع. فعموماً ترمز $t_{n-2; \gamma}$ إلى الرقم الذي يعبر عن الاحتمال بأن يكون متغير t بدرجات حرية $n-2$ أقل من ذلك الرقم يساوي γ .

*** مرة أخرى، نؤكد على ضرورة اختيار حجم الخطأ من النوع الأول قبل تحليل البيانات، فإذا لم يتم ذلك فسوف نقبل أي افتراض موضع الاهتمام عن طريق الحجم «الملائم» للخطأ من النوع الأول.

قد يكون من المفيد، أيضا، أن نشير إلى الشكل الذي يستعمله الاقتصاديون، عادة، عند تقرير نتائج الانحدار في بحوثهم أو دراساتهم. فإذا وقعت عينك على دورية اقتصادية أو كتاب يذكر دالة الاستهلاك التي قدرناها في الفصلين الثاني والثالث فقد تجد:

$$\hat{C} = 13 + 0.98Y_d \quad n = 10 \quad (3.23)$$

$$(5.6) \quad (0.01) \quad R^2 = 0.99$$

حيث الأرقام الموجودة بين القوسين تحت تقديرات المعلمات هي تقديرات للأخطاء المعيارية المناظرة (أي $\hat{\sigma}_b$ و $\hat{\sigma}_a$). ويمكن للقاري أن يبنى بهذه المعلومات فترات الثقة لمختلف المعاملات واختبار الفرضيات وهلم جرا.

نسبة t : قاعدة للحساب

على الرغم من أن الحجم الدقيق لفترة الثقة، ومن ثم، نتائج الاختبار، يعتمد على حجم العينة n، وعدد المعلمات التي ينبغي تقديرها، إلا أن الاقتصاديين يستخدمون عادة القواعد الحسابية العملية عند النظر إلى معادلة انحدار مقدرة. فمثلا إذا كانت قيمة المعلمة المقدرة أكبر من ضعف حجم الخطأ المعياري المقدر، يمكننا أن نستنتج في حالة الاختبار ذي الطرفين اختلاف تقدير المعلمة معنويا عن الصفر عند مستوى معنوية 5% أي إذا اعتبرنا فرضية العدم أن المعلمة تساوي الصفر مقابل الفرضية البديلة للاختبار ذي الدليلين، فإن هذه النتائج تؤدي إلى رفض فرضية العدم. أما إذا كان تقدير المعلمة أكثر من ثلاثة أمثال حجم الخطأ المعياري المقدر فسوف تكون هذه المعلمة، عموما، مختلفة عن الصفر عند مستوى معنوية 1%.

يمكن تبرير القواعد العملية للحساب لنسبة t هذه بسهولة. فمثلا إذا أردنا اختبار فرضية العدم $b = 0$ مقابل الفرضية البديلة $b \neq 0$ فإننا سوف نبني اختبارنا على أساس الفترة:

$$(\hat{b} \pm t_{n-2; 0.975} \hat{\sigma}_{\hat{b}}) \quad (3.24)$$

إذا أردنا أن يكون مستوى المعنوية عند 5%، فإذا لم تتضمن هذه الفترة الصفر، فسنرفض فرضية العدم. والآن يمكن أن تكون \hat{b} موجبة أو سالبة، ولكن $t_{n-2; 0.975}$

و $\hat{\sigma}_{\hat{b}}$ تكونان دائما موجبتين . وهنا لن تحتوي فترتنا للثقة على الصفر إذا كانت :

$$(\hat{b} - t_{n-2;0.975}\hat{\sigma}_{\hat{b}}) > 0 \quad \text{when } \hat{b} > 0, \quad (3.25)$$

أو

$$(\hat{b} + t_{n-2;0.975}\hat{\sigma}_{\hat{b}}) < 0 \quad \text{when } \hat{b} < 0$$

هذه الشروط يمكن إعادة كتابتها على النحو :

$$\frac{\hat{b}}{\hat{\sigma}_{\hat{b}}} > t_{n-2;0.975} \quad \text{when } \hat{b} > 0 \quad (3.26)$$

أو

$$\frac{\hat{b}}{\hat{\sigma}_{\hat{b}}} < t_{n-2;0.975} \quad \text{when } \hat{b} < 0$$

وأخيرا ، يمكن كتابة هذه الشروط بإيجاز على النحو :

$$\left| \frac{\hat{b}}{\hat{\sigma}_{\hat{b}}} \right| > t_{n-2;0.975} \quad (3.27)$$

لذلك إذا كانت القيمة المطلقة لـ $(\hat{b}_1 / \hat{\sigma}_{\hat{b}_1})$ تزيد على القيمة المعطاة بواسطة توزيع t ($t_{n-2;0.975}$) فسوف نرفض فرضية العدم $b=0$. وبمعنى آخر ، يمكننا اختبار الفرضية $b=0$ مقابل الفرضية البديلة $b \neq 0$ عند مستوي 5% من المعنوية عن طريق معرفة هل القيمة المطلقة للنسبة $\hat{b} / \hat{\sigma}_{\hat{b}}$ تزيد على $t_{n-2;0.975}$. وبملاحظة أن الاقتصاديين يتعاملون عادة مع عينات لاتقل في أحجامها عن 15 مشاهدة . فإذا اعتبرنا $n=15$ فإن $(t_{15-2;0.975} = t_{13;0.975}=2.16)$ ، أما إذا اعتبرنا $n = \infty$ فإن $(t_{\alpha;0.975}=1.96)$ ، وأخيرا فإن نظرة سريعة إلى جدول t تبين أن قيمة $t_j = 0.975$ وإذن ، فإن $\infty > j > 13$ تكون بين 1.96 و 2.16 . وهكذا تكون القاعدة على النحو «إذا زادت النسبة $\hat{b} / \hat{\sigma}_{\hat{b}}$ عن 2 بالقيمة المطلقة نرفض فرضية العدم بأن $b=0$ » .

ويطلق على نسبة مقدر المعلمة إلى خطئه المعياري : $\hat{b}/\hat{\sigma}_{\hat{b}}$ في الكتابات الإحصائية بنسبة t . وإذا أردنا تكوين الاختبارات للفروض عند مستوى معنوية 1% فإننا نجد $(t_{15-2,0.995} = t_{13,0.995} = 3.01)$ وأيضا $t_{\infty,995} = 2.58$ وفي هذه الحال يتطلب «كتقريب» أن تكون القيمة المطلقة لـ t أكبر من 3 قبل أن نصرح بأن تقدير المعلمة مختلف معنويا عن الصفر.

وبطريقة مشابهة جدا لما ذكر أعلاه، يمكننا اشتقاق نسبة t لاختبارات الفرضيات ذات الذيل الواحد. فعلى سبيل المثال، إذا أردنا اختبار الفرضية $b=0$ مقابل $b>0$ ، فسوف نرفض الفرضية $b=0$ عند 5% مستوى معنوية إذا :

$$(\hat{b} - t_{n-2,0.975} \hat{\sigma}_{\hat{b}}) > 0 \quad (3.28)$$

وبالمثل، إذا كانت الفرضية البديلة لـ $b=0$ هي $b<0$ فإننا سنرفض الفرضية $b=0$ إذا :

$$(\hat{b} + t_{n-2,0.975} \hat{\sigma}_{\hat{b}}) < 0 \quad (3.29)$$

والآن يمكننا إعادة كتابة كل من المعادلتين (3.28) و (3.29) على النحو :

$$\frac{\hat{b}}{\hat{\sigma}_{\hat{b}}} > t_{n-2,0.95} \quad (3.30)$$

و

$$\frac{\hat{b}}{\hat{\sigma}_{\hat{b}}} < -t_{n-2,0.95} \quad (3.31)$$

ومرة أخرى، فإن كل مانحتاجه هو تكوين نسبة t ومن ثم، مقارنتها بـ $t_{n-2,0.95}$ في المعادلة (3.30) والتي تناظر الفرضية البديلة ذات الطرف الواحد $b>0$ أو نقارنها بـ $t_{n-2,0.95}$ في (3.30) والتي تناظر الفرضية البديلة ذات الطرف الواحد $b>0$ أو نقارنها بـ $t_{n-2,0.95}$ كما في المعادلة (3.31) والتي تناظر الفرضية البديلة $b>0$. في هذه الحالة سيكون مدى القيم هو $t_{\infty,0.95} = 1.645$ و $t_{13,0.95} = 1.771$. وبالطبع يكون الشرط

الضروري لرفض فرضية العدم في أي من الحالتين هو :

$$\left| \frac{\hat{b}}{\hat{\sigma}_{\hat{b}}} \right| > t_{n-2;0.95} \quad (3.32)$$

وفي بعض الحالات، يقوم المؤلف بتوفير المشقة على القاريء من خلال قسمة قيمة المعامل بوساطة الخطأ المعياري المقدّر لتحديد نسبة t وإعطاء حاصل القسمة (قيمة نسبة t للعينه) بين قوسين أسفل المعامل. وهنا ينبغي أن نكون على حذر وأن نتأكد من الملاحظات التي يذكرها المؤلف لتوضيح ما إذا كانت الأرقام الموجودة بين الأقواس تمثل نسبة t أو تمثل الخطأ المعياري المقدّر. وعلى أي حال، ينبغي أن يكون واضحاً أن قواعد الحساب المرتبطة بنسب t تسهل كثيراً اختبار الفرضيات. ونكون قادرين، غالباً، على اختبار الفرضيات بدون الإشارة إلى جدول القيم لتوزيع t لأن نسب t تزيد، غالباً، على ثلاثة أو تكون أقل من واحد بالقيمة المطلقة.

(٣-٢) مشكلة شكل الدالة

لاشك أنك لاحظت خلال المناقشة في الفصلين الأول والثاني أننا افترضنا أن شكل العلاقة التي رغبنا في تقديرها هو الشكل الخطي، وبالتحديد فقد افترضنا أن :

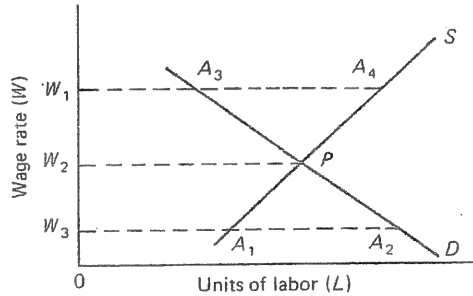
$$Y_t = a + bX_t + u_t$$

من الواضح أن هذا يعد شرطاً مقيداً جداً. وعادة ماتقترح النظرية الاقتصادية أو شكل انتشار النقط المشاهدة أن العلاقة بين المتغيرين غير خطية. والتساؤل الآن هو كيف نتعامل مع العلاقة غير الخطية بدلالة النموذج الخطي ؟

منحنى فليبس والتحويل العكسي

قد يكون من المفيد أن نعالج هذه المشكلة بدلالة علاقة اقتصادية فعلية :
منحنى فليبس Philips curve الذي يمثل حالة نموذجية للعلاقة غير الخطية بين النسبة

المثوية للتغير في الأجور (\dot{W}) ومعدل البطالة (R). اعتبر نمودجا لسوق العمل يكون مبسطا ومكونا من طلب وعرض. ويعتمد كل من الطلب على العمل وعرض العمل على معدل الأجر. يظهر هذا النموذج في الشكل رقم (٣-٣) حيث يحدد تقاطع الطلب والعرض الأجر التوازني W_2 . ولكن (من الناحية الأخرى) إذا كان معدل الأجر عند W_3 فإن الطلب على العمل سيزيد على العرض منه بالمقدار ($A_1 A_2$)، في هذه الحال، سنصرح بوجود فائض طلب موجب على العمل، وسيؤدي ذلك إلى ضغط معدلات الأجور إلى أعلى. وعلى العكس، إذا كانت $W=W_1$ فسيظهر فائض طلب سالب (أو فائض عرض) بالمقدار ($A_3 A_4$) مما يؤدي إلى ضغط مستويات الأجور لأسفل.



شكل رقم (٣-٣)

دعنا الآن نفترض آلية ديناميكية مبسطة للتعديل حتى نستطيع اكتشاف هذه العلاقات. وبالتحديد، نفترض أن القيمة المتوقعة لمعدل التغير في معدل الأجر (\dot{W}) من فترة لأخرى يرتبط نسبيا ومباشرة بمعدل فائض الطلب. ويبدو ذلك منطقيا حيث إنه كلما ازداد الفرق بين الطلب على العمال بواسطة رجال الأعمال وبين عرض العمل، فإننا نتوقع ضغطا على مستويات الأجور لأعلى. من أجل ذلك نفترض أن :

$$\dot{W}_t = \frac{(W_t - W_{t-1})}{W_{t-1}} = \alpha D_t^* + u_t \quad (3.33)$$

حيث إن $D_t^* = (D_t - S_t) / S_t$ هو معدل فائض الطلب في الفترة t و u_t هو الخطأ العشوائي.

لتقدير هذه العلاقة، نحتاج إلى مشاهدات عن \dot{W}_t و D_t^* . وعلى الرغم من توافر مشاهدات عن \dot{W}_t ، فعادة لا تتوفر مشاهدات (أو بيانات) عن D_t^* . فإذا اردنا أن يكون لدينا نموذج يمكن تطبيقه ويستطيع تفسير تعديلات الأجور فينبغي أن نجد متغيراً مرتبطاً بـ D_t^* حتى يمكن استخدامه مقياساً تقريبياً proxy. ومن المعقول أن نفترض أن معدل فائض الطلب D_t^* يرتبط بعلاقة منتظمة مع معدل البطالة R_t في الاقتصاد القومي، فإذا كانت البطالة منخفضة جداً كما هو الحال في سوق العمل المحدود من جانب العرض، فإننا نتوقع وجود فائض طلب كبير موجب والعكس صحيح. لذا لدينا سبب للاعتقاد بأن R_t و D_t^* يرتبطان ببعضهما بعضاً عكسياً، دعنا نفترض العلاقة التالية :

$$D_t^* = f(R_t) \quad (3.34)$$

ما الشكل الذي يجب أن تأخذه العلاقة (3.34)؟ أبسط الافتراضات هي أن العلاقة خطية على النحو :

$$D_t^* = e + gR_t \quad (3.35)$$

حيث إن e و g معلمات مع $g < 0$. ولكن، بقليل من التفكير، يتبين لنا أن المعادلة (3.35) ليس هو الشكل الأكثر ملائمة لهذه العلاقة. اعتبر الشكل (٣-٤) حيث يقاس معدل فائض الطلب على المحور الرأسي ومعدل البطالة على المحور الأفقي. عند النقطة 0 يكون لدينا فائض طلب يساوي الصفر، ويكون D_t^* موجبا أعلى 0 وسالبا أسفله. لاحظ أن فائض الطلب الصفري يناظر P في الشكل (٣-٣) حيث يتساوى الطلب وعرض العمل، ومن ثم، لاتجد ضغوطاً لتغيير مستوى الأجور. وتوجد، على أي الأحوال، بعض البطالة الاحتكاكية، بمعنى أنه، في ظل الاقتصاد

الحركي سيوجد بعض الأفراد في عملية انتقال من وظيفة لأخرى، ولكن، إذا كان فائض الطلب يساوي الصفر، فإن عدد الوظائف الشاغرة سيتساوى مع عدد الأفراد الذين يبحثون عن الوظائف. لذلك، فإن النقطة E في الشكل رقم (٣-٤) تمثل معدل البطالة الاحتكاكية (أي معدل البطالة الذي يناظر فائض طلب قدره الصفر ويناطر، أيضا، النقطة P في الشكل رقم (٣-٣)).

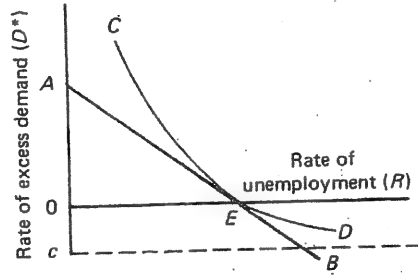
اعتبر بعد ذلك سلسلة من الفترات ذات الطلب الفائض (الفترات ذات القيم العالية لـ D^*). في هذه الحال ينبغي أن يقلل العدد المتزايد من الوظائف الشاغرة الوقت اللازم للحصول على الوظائف للعاطلين. لذلك نتوقع أن تصاحب القيم الأعلى لـ D^* قيمة أقل لـ R. غير أننا لا نتوقع أن يؤدي استمرار تزايد عدد الوظائف الشاغرة مع زيادة قيمة D^* إلى استمرار التناقص في R بكميات متساوية، وأحد اسباب ذلك هو أن معدل البطالة لا يمكن أن يأخذ قيمة سالبة. وكنتيجة لذلك لا يمكن أن تأخذ العلاقة بين D^* و R الشكل الخطي مثل AB في الشكل رقم (٣-٤). وكل هذا يعني أنه كلما أصبحت قيم D^* أكبر يكون الانخفاض المناظر في R أصغر. وهكذا ينبغي أن تكون العلاقة بين المتغيرين غير خطية مثل CD في الشكل رقم (٣-٤) حيث ينحني المنحنى إلى أعلى يسار النقطة E، وإلى يمين هذه النقطة، يكون لـ CD انحدار سالب أيضا، مشيرا إلى أن المعدلات الأعلى من R ترتبط بحالات فائض العرض. وللتوضيح، نفترض، أيضا، أن CD ينحني إلى أعلى يمين النقطة E.

إن أحد الأشكال الدالية التي «تقرب» منحنى مثل CD هو :

$$D_t^* = c + d \left(\frac{1}{R_t} \right), \quad \text{where } c < 0, d > 0 \quad (3.36)$$

حيث يفترض أن D^* تتغير عكسيا مع R. فإذا كانت d موجبة فإن هذا يعطينا منحنى ذا ميل سالب. إلا أنه منحنى غير خطي بمعنى أنه ينحني إلى أعلى مشيرا إلى أن معدل الانخفاض في R لكل وحدة إضافية في D^* يتناقص مع تزايد فائض

الطلب D^* . نفترض أن c سالبة، حيث تصبح D^* سالبة للقيم الأعلى من R . لدينا الآن مقياس تقريبي لـ D^* ، طالما توجد قيم مشاهدة لـ R_t ، فإذا ما عوضنا عن D_t^* من المعادلة (3.36) في معادلة الأجور نحصل على منحني فلييس المعروف:



شكل رقم (٣-٤)

$$\dot{W}_t = \alpha D_t^* + u_t = \alpha \left[c + d \left(\frac{1}{R_t} \right) \right] + u_t \quad (3.37)$$

$$= a + b \left(\frac{1}{R_t} \right) + u_t$$

حيث إن $a = \alpha c$ و $b = \alpha d$. وتوضح المعادلة (3.37) أن معدل التغير في الأجور يرتبط مباشرة مع مقلوب معدل البطالة. افترض أنه تتوافر لدينا عينة من القيم المشاهدة لكل من: \dot{W}_t و R_t . كيف يمكننا تقدير المعلمات a و b لهذه العلاقة غير الخطية؟ لاحظ أننا لن نحاول تقدير c أو d ولكننا سنحاول تقدير a و b لعلاقتنا المشاهدة في المعادلة (3.37).

على الرغم من أن العلاقة في المعادلة (3.37) هي علاقة غير خطية بين \dot{W}_t و R_t فإنه يمكن تفسيرها كعلاقة خطية بين \dot{W}_t ومقلوب R_t (أي $1/R_t$). لذلك يمكن تقدير المعادلة (3.37) باستخدام طرق تقدير النماذج الخطية السابق توضيحها إذا ما أدخلنا تغييرا طفيفا في الرموز. وبوضوح أكثر، دعنا نعرف متغيرا جديدا.

$$Z_t = \frac{1}{R_t} \quad (3.38)$$

حيث توجد لكل قيمة غير صفرية من R_t قيمة مناظرة من Z_t ، على سبيل المثال، في الجدول رقم (١-٣) إذا قمنا بإحلال Z_t محل $1/R_t$ في المعادلة (3.37) فإننا نحصل على :

$$\dot{W}_t = a + bZ_t + u_t \quad (3.39)$$

جدول رقم (١-٣) مصفوفة المشاهدات المفترضة

W	R	Z
0.02	0.06	16.7
0.04	0.04	25.0
0.05	0.03	33.3
.	.	.
.	.	.
.	.	.

بمعنى آخر، وبتحويل بسيط، فقد حولنا العلاقة غير الخطية في المعادلة (3.37) إلى علاقة خطية في المعادلة (3.39). وحيث، يمكننا استخدام نموذجنا الخطي للانحدار، وقيم \dot{W}_t و Z_t لتقدير المعلمات a و b . وعلى سبيل المثال، باستخدام الصيغ التي اشتققناها في الفصل الثاني، يكون لدينا :

$$\hat{b} = \frac{\Sigma(Z_t - \bar{Z})\dot{W}_t}{\Sigma(Z_t - \bar{Z})^2}, \quad (3.40)$$

$$\hat{a} = \bar{\dot{W}} - \hat{b}\bar{Z}.$$

ويتجميع النتائج السابقة، فإن مقدر القيمة المتوسطة لـ \hat{W}_t المناظر لقيمة معطاة R_t سيكون:

$$\hat{W}_t = \hat{a} + \hat{b} \left(\frac{1}{R_t} \right) \quad (3.41)$$

والآن يمكننا استخدام المعادلة (3.41) للتنبؤ. فعلى سبيل المثال، إذا كان معدل البطالة هو 5% فإننا نتوقع أن يكون معدل التغير في الأجور:

$$\hat{\dot{W}}_t = \hat{a} + \hat{b} \left(\frac{1}{0.05} \right) = \hat{a} + 20\hat{b} \quad (3.42)$$

ويعد التحويل العكسي reciprocal transformation أحد التحويلات التي يمكن استخدامها لتحويل العلاقة غير الخطية بين متغيرين إلى علاقة خطية. ولهذه التحويلات أهمية عظيمة لأنها تعني أن نموذج الانحدار الخطي البسيط الذي كوّناه ليس مقيدا ثقيلًا كبيرًا كما يظهر لأول وهلة. ومن خلال الاستخدام الحكيم لهذه التحويلات المختلفة، يصبح من الممكن أن نضع تشكيلة عريضة من العلاقات غير الخطية على شكل خطي، ويسمح لنا ذلك بتقدير معلمات مثل هذه النماذج باستخدام الطرق التي بنيناها فعلا. سندرس فيما يأتي تحويلين آخرين يستخدمان بكثرة في الاقتصاد القياسي.

التحويل اللوغاريتمي

افترض أننا نرغب في تقدير معلمات نموذج الإنتاج التالي:

$$Q_t = aL_t^b e^{u_t} \quad (3.43)$$

حيث:

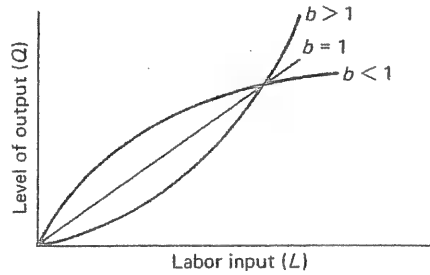
Q_t = حجم الإنتاج خلال الفترة t ،

L_t = مدخل العمل خلال الفترة t ،

e = حد ثابت يعادل بالتقريب 2.718،

u_t = الخطأ العشوائي في الفترة t و a و b هي المعلمات التي نرغب في تقديرها. في هذا النموذج، نفرض أن العمل هو عنصر الإنتاج الوحيد، وستخلى عن هذا الافتراض في الأجزاء الأخيرة من هذا الكتاب.

وعلى افتراض أن المعادلة (3.43) تمثل وصفا دقيقا لدالة الإنتاج فقد نهتم بخاصة بالمعلمة b ، أو أنه قد تكون لدينا افتراضات معينة حولها، لأن قيمة المعلمة b تبين لنا ما إذا كان هناك تناقص أو، ثبات أو تزايد في غلة الحجم، وتناظر هذه الحالات $b < 1$ و $b = 1$ أو $b > 1$ على الترتيب كما سيتضح من الشكل رقم (٣-٥).



شكل رقم (٣-٥)

وأولى المشاكل القياسية في المعادلة (3.43) هي أن العلاقة غير خطية ومانحتاجه، مرة أخرى، طريقة تحويلها إلى علاقة خطية حتى يمكننا تطبيق طرق التقدير التي عرفناها واستخدمناها من قبل.

والتحويل الذي نبحث عنه هو التحويل اللوغاريتمي. فمثلا، إذا أخذنا اللوغاريتمات لكل جانب من المعادلة (3.43) يكون لدينا:

$$\ln Q_t = \ln a + b \ln L_t + u_t \quad (3.44)$$

حيث ترمز \ln إلى «اللوغاريتم الطبيعي» للمتغير (أي اللوغاريتم الذي يكون أساسه e التي تعادل 2.718 بالتقريب). يتضح لنا الآن أن المعادلة (3.44) خطية في لوغاريتمات المتغيرات، ولذا نقوم بالتحويل اللوغاريتمي التالي :

دع :

$$Q_t^* = \ln Q_t, \quad a^* = \ln a, \quad \text{و} \quad L_t^* = \ln L_t \quad (3.45)$$

وباستخدام هذه التعويضات في المعادلة (3.44)، يصبح لدينا :

$$Q_t^* = a^* + bL_t^* + u_t \quad (3.46)$$

والآن ينبغي أن يتضح أنه إذا عملنا الافتراضات المعتادة المرتبطة بالخطأ العشوائي u_t ، فإنه يمكن تقدير المعلمات a و b في المعادلة (3.46) بوساطة طريقة المتغير المساعد. وبالتحديد، فعن طريق أخذ اللوغارتمات الطبيعية للقيم المشاهدة لـ Q_t و L_t يمكن حساب

$$\hat{b} = \frac{\Sigma(L_t^* - \bar{L}^*)Q_t^*}{\Sigma(L_t^* - \bar{L}^*)^2} \equiv \frac{\Sigma(\ln L_t - \ln \bar{L})L_n Q_t}{\Sigma(\ln L_t - \ln \bar{L})^2} \quad (3.47)$$

$$\hat{a}^* = \bar{Q}^* - \hat{b}\bar{L}^* \equiv \ln \bar{Q} - \hat{b}\ln \bar{L}$$

حيث

$$\bar{L}^* = \Sigma L_t^* / n, \quad \text{و} \quad \bar{Q}^* = \Sigma Q_t^* / n$$

ولما كنا نعرف أن مقدراتنا غير متحيزة، فسيكون لدينا

$$E(\hat{b}) = b \quad \text{و} \quad E(\hat{a}^*) = a^* \quad (3.48)$$

تعطينا الصيغة في المعادلة (3.47) مقدرًا غير متحيز لـ a متغير مدخل العمل في المعادلة (3.43). وباستخدام هذا المقدر يمكننا أن نتبين حقيقة ما إذا كانت نتائجنا تدل على وجود ظاهرة تناقص الغلة بالنسبة لعنصر العمل.

نلاحظ، أيضًا، أن طريقتنا في التقدير تعطي مقدرًا غير متحيز لـ a^* ، ولكننا في الحقيقة نهتم بقيم المعلمة a وليست a^* ، طالما أن a هي المعلمة التي تظهر في دالة الإنتاج. وطالما أن $a^* = \ln a$ فإننا نأخذ عكس اللوغارتم ولذا تصبح $a = e^{a^*}$ ، ويكون المقدر المقترح لـ a هو:

$$\hat{a} = e^{\hat{a}^*} \quad (3.49)$$

ولكن \hat{a} ليس مقدرًا غير متحيز لـ a على الرغم من أن $E(\hat{a}^*) = a^*$ أي أن

$E(\hat{a}) \neq e^E(\hat{a}^*) = e^{a^*} = a$. لعنا نتذكر أننا اشرنا في ملحق الفصل الأول إلى أن القيمة المتوقعة لدالة غير خطية مثل $E(\hat{e}^*)$ لا تتساوي عموماً مع دالة القيمة المتوقعة $\left[e^{E(\hat{a}^*)} \right]$ ، وهذا مثال لهذه القاعدة . ولحسن الحظ فإنه لا يزال يمكننا إثبات أن \hat{a} مقدر متسق لـ a . خلاصة القول أنه إذا كان لدينا دالة ذات الشكل العام :

$$Y_t = aX_t^b e^{u_t} \quad (3.50)$$

فإنه يمكننا استخدام التحويل اللوغاريتمي لوضع تلك الدالة في الشكل الخطي :

$$Y_t^* = a^* + bX_t^* + u_t \quad (3.51)$$

حيث تعني (*) اللوغارتم الطبيعي للمتغير المناظر . ويمكننا تقدير المعادلة (3.51) باستخدام طرق التقدير الخطية التي يمكن استخدامها للحصول على تقدير غير متحيز لـ b ، وبأخذ عكس اللوغارتمات نحصل على تقدير متحيز ولكن في الأقل ، متسق لـ a .

ويعد الشكل اللوغاريتمي من الاشكال الدالية الشائعة جداً للنماذج الاقتصادية بسبب أنه يمكن تفسير معامل الانحدار على أنه مرونة المتغير التابع للمتغير المستقل . فعلى سبيل المثال في المعادلة (3.51) أو (3.50) تكون مرونة القيمة المتوقعة لـ Y_t للمتغير X_t في الحقيقة b^* لذلك يتضمن النموذج (3.51) ثبات المرونة .

التحويل شبه اللوغاريتمي The semilog transformation

تظهر أهمية التحويل شبه اللوغاريتمي في تكوين النماذج التي تحتوي على

* على سبيل المثال يتضح لنا من المعادلة (3.50) أن القيمة المتوقعة لـ Y_t (عند قيمة معينة لـ X_t) هي $Y_t^m = aX_t^b E(e^{u_t})$. ويتضمن الافتراض بأن X_t مستقلان أن يكون $E(e^{u_t})$ مساوياً لرقم ثابتا (مثلاً، C) لن يكون مساوياً (عموماً) الوحدة مثلاً، $\left(E(e^{u_t}) \neq e^{E(u_t)} = e^0 = 1 \right)$ وقد نبر عن القيمة المتوقعة لـ Y_t^m المناظرة لـ X_t على النحو $Y_t^m = a_1 X_t^b$ ، حيث $a_1 = ac$. بأخذ التفاضل بعد ذلك لـ Y_t^m بالنسبة لـ X_t نحصل على $\frac{dY_t^m}{dX_t} = a_1 b X_t^{b-1} = \frac{bY_t^m}{X_t}$

وبحل هذه المعادلة بهدف الحصول على b ، نجد أن :

$$b = \frac{dY_t^m / Y_t^m}{dX_t / X_t}$$

وهي مرونة Y_t^m بالنسبة لـ X_t .

معدلات للنمو. افترض، على سبيل المثال، أننا نحتاج إلى تقدير متوسط المعدل السنوي للزيادة في حجم قوة العمل في الولايات المتحدة الأمريكية على مدى فترة معينة. فقد نعتقد أنه، خلال تلك الفترة، زادت قوة العمل بمعدل سنوي ثابت مع تغيرات طفيفة ناتجة عن الحوادث العشوائية المختلفة. فإذا كان ذلك صحيحاً فقد يكون من الممكن افتراض علاقة مثل :

$$L_t = a(1+g)^t e^{u_t}, \quad t=1,2,\dots,n \quad (3.52)$$

حيث :

L_t = حجم قوة العمل خلال السنة t .

a = معلمة

g = معلمة وهي معدل النمو المركب في L_t

u_t = الخطأ العشوائي.

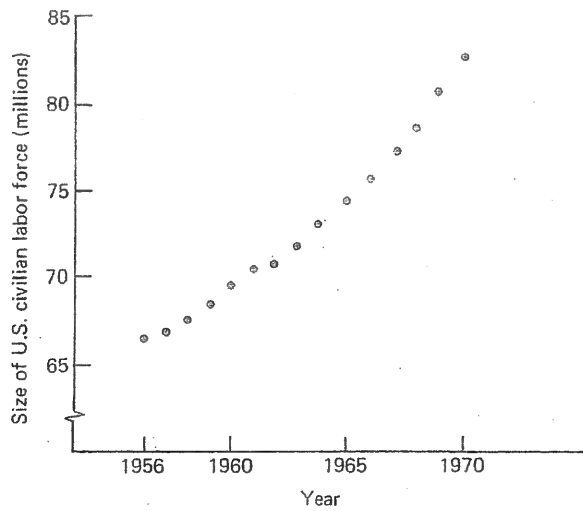
يلاحظ أن المتغير المستقل t في المعادلة (3.52) يظهر بوصفه قوة (أس) في المعادلة (3.50) بينما يظهر المتغير المستقل X_t (على نحو مغاير) مرفوعاً لقوة ثابتة b . ولكن كلا من المعادلة (3.52) والمعادلة (3.50) يظهر في شكل حاصل ضرب. وهكذا، وكما هو متوقع، سنأخذ اللوغاريتمات لتحويل العلاقة إلى الشكل الخطي. وقبل أن نفعل ذلك، ينبغي أن نشير إلى إحدى خصائص مسارات النمو الزمنية. نعرض في جدول (٣-٢) بيانات عن حجم قوة العمل المدنية في الولايات المتحدة الأمريكية خلال السنوات ١٩٥٦-١٩٧٠ م. ويرسم هذه البيانات كما يظهر في الشكل (٣-٦)، نجد أن المنحنى يصبح أكثر انحداراً بمرور الوقت مما يوحي بأن قوة العمل قد نمت بمعدلات أكبر في السنوات الحديثة. إلا أن هذا في الحقيقة توهم فحسب لأن معدلاً معيناً للنمو سوف يولد زيادات مطلقة لقوة العمل سنة بعد أخرى، بمعنى أن الأساس الذي يبنى عليه النمو سيكون أعلى في السنوات الحديثة منه في السنوات السابقة (فمثلاً تكون 3% من X أكبر من 3% من Y إذا كانت $Y < X$).

جدول رقم (٣-٢) قوة العمل المدنية بالولايات المتحدة بملايين الأفراد

السنة	قوة العمل
١٩٥٦	٦٦,٦
١٩٥٧	٦٦,٩
١٩٥٨	٦٧,٦
١٩٥٩	٦٨,٤
١٩٦٠	٦٩,٦
١٩٦١	٧٠,٥
١٩٦٢	٧٠,٦
١٩٦٣	٧١,٦
١٩٦٤	٧٣,١
١٩٦٥	٧٤,٥
١٩٦٦	٧٥,٤
١٩٦٧	٧٧,٣
١٩٦٨	٧٨,٧
١٩٦٩	٨٠,٧
١٩٧٠	٨٢,٧

المصدر: التقرير الاقتصادي للرئيس، واشنطن، الولايات المتحدة، مكتب الطباعة الأمريكي،

فبراير ١٩٧١م، ص ٢٢٢.



شكل رقم (٣-٦)

فإذا أخذنا الآن اللوغاريتمات لطرفي المعادلة (3.52) نحصل على :

$$\ln L = \ln a + t \ln(1 + g) + u_t \quad (3.53)$$

فإذا جعلنا

$$L_t^* = \ln L_t$$

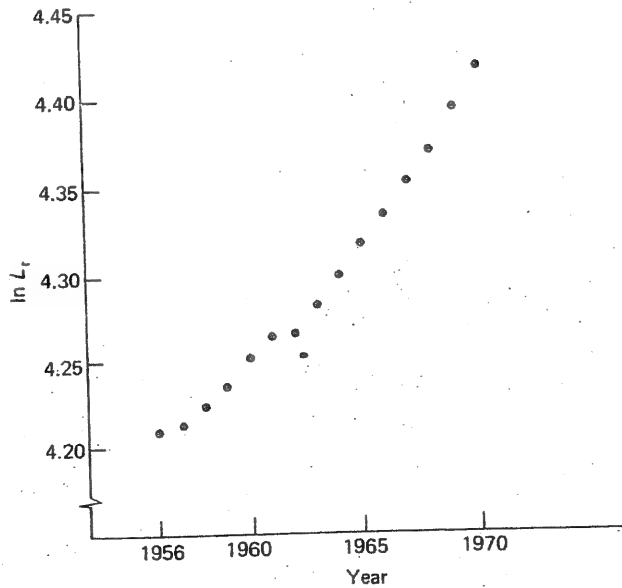
$$a^* = \ln a, \quad (3.54)$$

$$b^* = \ln(1 + g)$$

فإننا نحصل على

$$L_t^* = a^* + b^* t + u_t \quad (3.55)$$

وتوضح لنا المعادلة (3.55) أن معدل النمو المركب يتضمن علاقة خطية بين $\ln L_t$ و t وليس بين L_t و t . فإذا رسمنا قيم لوغاريتم L_t بدلا من قيم L_t عبر الزمن لشاهدنا النقاط تقارب المسار الخطي (كما يظهر في الشكل ٧-٣).



شكل (٧-٣)

ولتقدير المعلمات الموجودة في المعادلة (3.55) أي (a^* و b^*) ينبغي أن تتوافر لدينا مشاهدات حول L_t و t لكل سنة من السنوات التي ندرسها. وللتوضيح، دعنا نعود إلى القيم المشاهدة لقوة العمل بالولايات المتحدة الأمريكية خلال السنوات ١٩٥٦-١٩٧٠ م. وتزودنا القيم المشاهدة لـ L_t لكل من هذه السنوات مباشرة بالقيم المشاهدة المناظرة للوغاريتم L_t . يحصل على مشاهدات t ، بسهولة، عن طريق وضع أرقام للسنوات على التتابع [أي ١٩٥٦ السنة الأولى ($t=1$)، و ١٩٧٠ السنة الخامسة عشرة ($t=15$)] ويوضح الجدول (٣-٣) ذلك.

جدول (٣-٣) قوة العمل المدنية بالولايات المتحدة (بالملايين)

السنة	t	L	$\ln L$
١٩٥٦	١	٦٦,٦	٤,١٩٩
١٩٥٧	٢	٦٦,٩	٤,٢٠٣
١٩٥٨	٣	٦٧,٦	٤,٢١٤
١٩٥٩	٤	٦٨,٤	٤,٢٢٥
١٩٦٠	٥	٦٩,٦	٤,٢٤٣
١٩٦١	٦	٧٠,٥	٤,٢٥٦
١٩٦٢	٧	٧٠,٦	٤,٢٥٧
١٩٦٣	٨	٧١,٨	٤,٢٧٤
١٩٦٤	٩	٧٣,١	٤,٢٩٢
١٩٦٥	١٠	٧٤,٥	٤,٣١١
١٩٦٦	١١	٧٥,٨	٤,٣٢٨
١٩٦٧	١٢	٧٧,٣	٤,٣٤٨
١٩٦٨	١٣	٧٨,٧	٤,٣٦٦
١٩٦٩	١٤	٨٠,٧	٤,٣٩١
١٩٧٠	١٥	٨٢,٧	٤,٤١٥

ويمكننا، بسهولة، من المعلومات الواردة في الجدول (٣-٣) حساب:

$$\hat{b}^* = \frac{\sum(t - \bar{t})L_t^*}{\sum(t - \bar{t})^2} = 0.0153$$

$$\hat{a}^* = \bar{L}_t^* - \hat{b}^* \bar{t} = 4.17$$

ولما كانت $[\ln(1+g) = b^*]$ ، نجد أن $(g = e^{b^*} - 1)$. لذا، نقدر معدل النمو (g) عن طريق $(g = e^{b^*} - 1 = 2.718^{(0.0153)} - 1 = 0.016)$. وهكذا، يصبح تقديرنا لمعدل النمو السنوي لقوة العمل ١,٦٪. ولأسباب ذكرت في المبحث السابق، يكون مقدرنا لـ $a = e^{a^*}$ متحيزاً إلا أنه متسق.

هناك طريقة أخرى لحساب معدلات النمو تظهر، أحياناً، في الكتابات الإحصائية ويمكن توضيحها على النحو التالي: خذ القيمة الأولية لقوة العمل L_t (٦٦,٦ مليوناً عام ١٩٥٠م) والقيمة المشاهدة الأخيرة (٢٢,٧ مليوناً عام ١٩٧٠م). وبعدها وبمساعدة جدول اللوغاريتمات احسب متوسط معدل النمو السنوي (أي حدد معدل النمو السنوي الذي يؤدي إلى جعل ٦٦,٦ تصبح ٨٢,٧ بعد ١٥ عاماً).

هذه الطريقة البديلة لانصح باستخدامها، حيث إنها تأخذ النقطتين الأولى والأخيرة فقط في الشكل (٧-٣) (رسم $\ln L_t$ مقابل t) وإيجاد ميل المنحنى الذي يربط بين هاتين النقطتين. وبمعنى آخر أنه يجعل الخط يمر بنقطتين، فقط، من نقاط شكل الانتشار. تتشابه هذه الطريقة مع الطريقة الأولى، فقط، إذا افترضنا أن العينة المتاحة ذات حجم $n=2$ ، ولكن، لجميع العينات من الحجم $(n>2)$ تكون هذه الطريقة البديلة أقل دقة من الطريقة الأولى بسبب كل المعلومات التي تتجاهلها.

استخدام التحويلات: تعميمات

وجدنا في حالات عديدة أن افتراضنا المسبق قد يوحي بشكل معين عن العلاقة غير الخطية بين المتغيرات موضع الاهتمام. وغالباً ما سنجد التحويل الملائم لهذه العلاقة إلى الشكل الخطي ثم نطبق بعد ذلك طرق التقدير الخطية. والآن ينبغي أن تكون التحويلات التي تقوم بهذه المهمة واضحة. فمثلاً، إذا كان النموذج هو:

$$Y_t = a + bf(X_t) + u_t \quad (3.56)$$

حيث إن $f(X)$ هي دالة* ما في X_t ، يمكننا تحديد متغير $Z_t = f(X_t)$ ولذا يكون لدينا نموذج خطي يربط بين Y_t و Z_t . وتعميم العلاقة في المعادلة (3.56) هو:

$$g(Y_t) = a + bf(X_t) + u_t \quad (3.57)$$

حيث إن f و g هما دالتان ربما (غير خطيتين) لـ Y_t و X_t . في هذه الحالة، نكون لدينا علاقة خطية بين Z_{1t} و Z_{2t} حيث إن $[Z_{1t} = g(Y_t)]$ و $[Z_{2t} = f(X_t)]$. وأخيراً فإن نموذجاً يأخذ شكل حاصل الضرب من النوع:

$$g(Y_t) = af(X_t)^\alpha e^{u_t} \quad (3.58)$$

سيكون نموذجاً خطياً في Z_{1t}^* و Z_{2t}^* ، حيث إن $Z_{1t}^* = \ln g(Y_t)$ ، $Z_{2t}^* = \ln f(X_t)$. وعلى سبيل تدريب للقارئ، عليك أن تثبت أن النموذج:

$$\frac{1}{Y_t} = aX_t^{2\alpha} e^{u_t} \quad (3.59)$$

هو نموذج خطي في كل من $\ln X_t^*$ ، $\ln(1/Y_t)$ **.

في عديد من الحالات تمدنا النظرية الاقتصادية بقليل من المساعدة لتحديد الشكل الدقيق للعلاقة موضع الاهتمام. فقد تقترح النظرية (مثلاً) أن الكمية المطلوبة من سلعة ما تتغير عكسياً مع السعر. ولكنها لا تخبرنا عن شكل هذه العلاقة، هل هي علاقة خطية أو لوغاريتمية أم شكل أكثر تعقيداً. في مثل هذه الحالات، يمكن، أحياناً، تحديد الشكل الدالي للنموذج عن طريق الفحص البسيط لشكل نقاط الانتشار. أي أن الشكل الدالي يختار ليتوافق مع شكل الانتشار. لكن هذا المنهج (منهج شكل الانتشار) قد يكون مفيداً، فقط،

* يفترض أن $f(X_t)$ لا تحتوي على معلمة غير معلومة. أي أنه، إذا كانت لدينا مشاهدات عن X_t ، فإنه

يمكننا أن نكون مشاهدات عن $f(X_t)$.

** لاحظ أنه يمكن كتابة $X^{2\alpha}$ على النحو $(X^2)^\alpha$.

عندما يكون هناك متغيران في العلاقة.* ولما كانت معظم النماذج الاقتصادية تحتوي على عدد كبير من المتغيرات، فسوف نؤجل المناقشة الأكثر عمقا في هذا الموضوع حتى الفصل الخامس.

(٣-٣) الترجيح ووحدة القياس

من المهم عند إجراء الحسابات الفعلية لمعادلة الانحدار استخدام وحدات قياس معقولة ذلك من أجل تبسيط الحسابات في بعض الحالات وتسهيل تفسير النتائج في حالات أخرى. اعتبر، على سبيل المثال، الجدول (٣-٤) للقيم المشاهدة للاستهلاك الكلي وللدخل المتاح.

جدول رقم (٣-٤)

السنة	الدخل المتاح بالدولارات	الإنفاق الاستهلاكي بالدولارات
١٩٦٠	٣٦٤,٠٠٠,٠٠٠,٠٠٠	٣٢٥,٠٠٠,٠٠٠,٠٠٠
١٩٦١	٣٥٠,٠٠٠,٠٠٠,٠٠٠	٣٢٥,٠٠٠,٠٠٠,٠٠٠
⋮	⋮	⋮
١٩٦٩	٦٣٠,٠٠٠,٠٠٠,٠٠٠	٥٧٦,٠٠٠,٠٠٠,٠٠٠

افترض أننا نريد استخدام هذه القيم المشاهدة لـ C_t و Y_{dt} لتقدير دالة خطية للاستهلاك:

$$C_t = a + bY_{dt} + u_t$$

* توجد مشكلة أخرى أكثر صعوبة. إذا حُلِدَ شكل الدالة بوساطة الفحص الأولى للبيانات، وتمثل في وجود عنصر دائرية circularity. لذا ينبغي علينا، نظريا، أن نحدد أولا نموذجنا ثم نختبره بعد ذلك في ضوء البيانات. أما إذا حدد شكل العلاقة عن طريق الفحص الأولى للبيانات، فإن المنهج الصحيح سيكون استخدام ذلك الشكل ثم اختبار النموذج بعد ذلك مع مجموعة جديدة من البيانات، ولما كان الاقتصاديون لا يجدون، عادة، أكثر من عينة واحدة فإن عنصر الدائرية هذا، ولسوء الحظ، يغفل غالبا.

هل من الضروري الاحتفاظ بالأصفار التسعة إلى يمين كل مشاهدة لـ C_t و Y_{dt} ؟
الاجابة هي لا ، إذ يمكننا أن نقلل العبء على أنفسنا إذا أسقطنا هذه الأصفار عن
طريق قياس كل متغير بالوحدات الملائمة (في حالتنا هذه بـبلايين الدولارات بدلا
من الدولارات). وقياسا على ذلك ، نجد الفلكيين ، مثلا ، يقيسون المسافة بالسنوات
الضوئية وليس بالبوصات.

فإذا قسنا متغيراتنا ببلايين الدولارات فإن بياناتنا في الجدول رقم (٣-٤)
تصبح هي البيانات الموجودة في الجدول (٣-٤أ) ويمكننا بسهولة استخدام البيانات
الموجودة في هذا الجدول في تقدير المعلمات a و b لدالة الاستهلاك.

جدول رقم (٣-٤أ)

السنة	الدخل المتاح ببلايين الدولارات	الإنفاق الاستهلاكي ببلايين الدولارات
١٩٦٠	٣٥٠	٣٢٥
١٩٦١	٣٦٤	٣٢٥
⋮	⋮	⋮
١٩٦٩	٦٣٠	٥٧٦

والآن افترض أن باحثا آخر يقيس الدخل المتاح والإنفاق الاستهلاكي بمئات
البلايين من الدولارات. ويظهر جدول له للقيم الملاحظة كما في الجدول (٣-٤ب).
باستخدام هذه البيانات ، يمكن أيضا ، للباحث الثاني تقدير المعلمات a و b لدالة
الاستهلاك. والآن نسأل عن العلاقة بين تقديرات المعلمات المبنية على الجدول
(٣-٤أ) وتلك المؤسسة على الجدول (٣-٤ب).

جدول رقم (٣-٤) ب.

السنة	الدخل المتاح بمئات البلايين من الدولارات	الإنفاق الاستهلاكي بمئات البلايين من الدولارات
١٩٦٠	٣,٥٠	٣,٢٥
١٩٦١	٣,٦٤	٣,٢٥
⋮	⋮	⋮
١٩٦٩	٦,٣٠	٥,٧٦

وتشابه العلاقة بين تقديرات المبيعات المبنية على الوحدات المختلفة للقياس، بالضبط، تلك العلاقة الموجودة بين القيم المناظرة للمبيعات. افترض (مثلاً، أننا نقيس الدخل المتاح والإنفاق الاستهلاكي ببلايين الدولارات ونعبر عن النموذج على النحو:

$$c_t = a + by_{dt} + u_t \quad (3.60)$$

تمثل المعلمة a (في هذا النموذج) كمية الانفاق الاستهلاكي (ببلايين الدولارات) عندما يكون الدخل المتاح مساوياً للصفر. بينما تعبر المعلمة b عن الميل الحدي للاستهلاك. افترض الآن أننا نقسم كل حد في المعادلة (3.60) على مائة، إذن سنحصل على:

$$C_t = A + bY_{dt} + U_t \quad (3.61)$$

حيث إن:

$$C_t = \left(\frac{1}{100}\right)c_t, \quad A = \left(\frac{1}{100}\right)a, \quad Y_{dt} = \left(\frac{1}{100}\right)y_{dt}, \quad U_t = \left(\frac{1}{100}\right)u_t,$$

المعادلة (3.61) هي دالة استهلاك تربط بين الإنفاق الاستهلاكي والدخل المتاح عندما تقاس هذه المتغيرات بمئات البلايين من الدولارات. واشتقت هذه المعادلة من المعادلة (3.60)، ولذا، يجب أن تكون متناسقة معها. فمثلاً، إذا كانت $(Y_{dt} = 0, C_t = A + u_t)$ أو بالضرب في مائة $(c_t = a + u_t)$. فإن الباحث الذي يستخدم البيانات الواردة في الجدول (3.4A) يعتبر في الحقيقة النموذج (3.60).

تتضمن هذه العبارة أنه بمقارنة المعادلة (3.60) بالمعادلة (3.61) ينتهي الباحثون إلى التقدير نفسه للميل الحدي للاستهلاك، بينما سيحصل الباحث المستخدم للجدول (٣-٤أ) على حد ثابت أكبر مائة مرة عن ذلك الحد الثابت المشتق عن الجدول رقم (٣-٤ب). وكما رأينا فإن التقديرات لن تكون غير متناسقة لأن المتغيرات معرفة بطريقة مختلفة.

وإثبات صحة هذه العلاقات من السهولة بمكان. لاحظ أولاً أن

$$\bar{Y}_d = (1/100)y_d \text{ وأن } \bar{C} = (1/100)c \text{ . حيث}$$

$$\hat{b}_1 = \frac{\Sigma(Y_{dt} - \bar{Y}_d)C_t}{\Sigma(Y_{dt} - \bar{Y}_d)^2} = \frac{\Sigma(y_{dt} - \bar{y}_d)c_t}{\Sigma(y_{dt} - \bar{y}_d)^2} = \hat{b}_2 \quad (3.62)$$

حيث إن b_1 سيكون المقدّر لـ b الذي يحصل عليه بواسطة الجدول رقم (٣-٤ب) والنموذج (3.61)، و b_2 يناظر الجدول رقم (٣-٤أ) والنموذج (3.60) وبالنسبة للحدود الثابتة يكون لدينا:

$$\hat{A} = \bar{C} - \hat{b}\bar{Y}_d = \frac{1}{100}(\bar{c} - \hat{b}\bar{y}_d) = \frac{1}{100}\hat{a} \quad (3.63)$$

دعنا الآن نعمّم نتائجنا، اعتبر النموذج:

$$y_t = a + bx_t + u_t \quad (3.64)$$

والآن دع:

$$Y_t = s_1 y_t, \quad X_t = s_2 x_t \quad (3.65)$$

حيث إن s_1 و s_2 ثوابت أو عوامل ترجيح scale factors وبإحلال المعادلة (3.65) محل المعادلة (3.64) تكون لدينا العلاقة التالية بين Y_t و X_t :

$$\begin{aligned} Y_t &= as_1 + \left(\frac{bs_1}{s_2} \right) X_t + s_1 u_t \\ &= A + BX_t + U_t \end{aligned} \quad (3.66)$$

حيث إن $U_i = s_1 u_i$ و $B = bs_1/s_2$ ، $A = s_1 a$ ، المعادلة (3.64)، بينما قام باحث آخر باختيار وحدات ترجيح أخرى (فأسقط، مثلاً، الأصفار، غير الضرورية) لـ Y_i و X_i كما في المعادلة (3.65) ثم قام بتقدير المعادلة (3.66) بعد ذلك فإن العلاقة بين تقديرات معلماتها المناظرة يمكن الحصول عليها بوساطة:

$$\hat{A} = s_1 \hat{a}, \quad \hat{B} = \hat{b} \left(\frac{s_1}{s_2} \right) \quad (3.67)$$

في ضوء المعادلة (3.67) تكون العلاقات بين تباينات المقدرات على النحو:

$$\sigma_{\hat{A}}^2 = s_1^2 \sigma_{\hat{a}}^2, \quad \sigma_{\hat{B}}^2 = \left(\frac{s_1}{s_2} \right)^2 \sigma_{\hat{b}}^2 \quad (3.68)$$

وعلى الرغم من أننا لن نقوم بإثباتها هنا، فإنه يمكن إثبات أن العلاقات بين مقدرات التباينات هي نفسها، بالضبط، كنظائرها في المعادلة (3.68):

$$\hat{\sigma}_{\hat{A}}^2 = s_1^2 \hat{\sigma}_{\hat{a}}^2, \quad \hat{\sigma}_{\hat{B}}^2 = \left(\frac{s_1}{s_2} \right)^2 \hat{\sigma}_{\hat{b}}^2 \quad (3.69)$$

فإذا اعطينا عوامل الترجيح، فإنه يمكن، عند ذلك، اشتقاق نتائج إحدى هذه الدراسات مباشرة من النتائج الأخرى.

قبل أن نعطي مثالا لنظرية الترجيح هذه، ربما ينبغي أن نشير إلى ما قد يكون واضحاً. إذا كانت لدينا افتراضات خاصة بالمعلمات a أو b في المعادلة (3.64) فإن هذه الافتراضات قد تختبر أما بدلالة المعادلة (3.64) أو المعادلة (3.66). فمثلاً، نجد الافتراض $b = b^0$ يماثل الافتراض $B = b^0(s_1/s_2)$. ويمكن أن نذكر عبارات مشابهة للافتراضات المختصة بالمعلمة a . على الرغم من أننا لن نقوم بذلك هنا، فإنه يمكن إثبات أن قبول افتراض معين أو رفضه لـ a أو b والذي يختار بدلالة \hat{a} أو \hat{b} كما يشتق من المعادلة (3.64) فقط إذا كان الافتراض المناظر والمربط

بـ A أو B والذي يختار في المعادلة (3.66) قد قبل أو رفض. * وبمعنى آخر، فإن نسب t لـ \hat{a} و \hat{b} تكون متماثلة مع تلك لـ \hat{A} و \hat{B} .

مثال

دعنا الآن نأخذ تطبيقاً مبسطاً لمبادئ الترجيح هذه. افترض أننا مهتمون، مرة أخرى، بدالة الاستهلاك:

$$c_t = a + by_{dt} + u_t \quad (3.70)$$

افترض، أيضاً أننا نجمع بيانات عن y_{dt} ، c_t ، وأتينا نقدر معلماتنا a و b ، وأخيراً نختبر الافتراض $b = b_0$. افترض أننا قد علمنا أن البيانات التي استخدمناها غير دقيقة، وبالتحديد بسبب الطريقة التي جمعت بها البيانات، فإن أرقامنا عن الدخل المتاح تفوق البيانات بمقدار ١٠٪، ولكن، يفترض أن بياناتنا المرتبطة بالاستهلاك دقيقة. ويصبح التساؤل حول ما إذا كان من المطلوب إعادة الدراسة بالكامل باستخدام البيانات المصححة.

دع y_{dt}^* هو مقياسنا للدخل المتاح. يتضمن خطأ القياس أن:

$$y_{dt}^* = (1.1)y_{dt} \quad (3.71)$$

حيث إن y_{dt} هي القيمة الحقيقية للدخل المتاح. وبإحلال المعادلة (3.71) في دالتنا للاستهلاك بالمعادلة (3.70) نحصل على:

$$c_t = a + \left(\frac{b}{1.1} \right) y_{dt}^* + u_t \quad (3.72)$$

$$= a + By_{dt}^* + u_t$$

* يمكن للقارئ أن يقطع نفسه بهذا عن طريق ملاحظة أنه، في ضوء المعادلة (3.69)

$$\frac{\hat{B} - B}{\hat{\sigma}_{\hat{B}}} \equiv \frac{\hat{b} - b}{\hat{\sigma}_{\hat{b}}} \quad \text{و} \quad \frac{\hat{A} - A}{\hat{\sigma}_{\hat{A}}} \equiv \frac{\hat{a} - a}{\hat{\sigma}_{\hat{a}}}$$

وهكذا، فإن فترات الثقة لـ A و B المبنية على المعادلة (3.66) هي، ببساطة، مرجحة لأعلى أو أسفل لفترات المناظرة كما تشتق من المعادلة (3.64).

حيث إن $B = (b/1.1)$.

المعادلة (3.72) هي النموذج المرتبط باستخدام بياناتنا غير الدقيقة. باستخدام نتائجنا بالمعادلة (3.67) يمكننا أن نرى من المعادلة (3.72) أن تقديرنا للحد الثابت a بالإضافة إلى نتائج اختبار الفرضية المرتبطة بذلك الحد لا تزال صحيحة. ولكن تقديرنا للميل الحدي للاستهلاك قد يكون منخفضا جدا لأن مقدرا غير متحيز لـ MPC يمكن أن يكون $\hat{b} = \hat{B}(1.1)$. وهكذا، ينبغي أن أن نضرب تقديرنا في (1.1) إضافة إلى ذلك، ينبغي علينا أن نعيد اختبار فرضيتنا المرتبطة بقيمة b . ومن السهل القيام بذلك إذا لاحظنا أن الفرضية (مثلاً $b = b_0$) تتضمن الفرضية $B = b_0/1.1$. وفي الحقيقة فإن قدرا ضئيلا من عملنا الأصلي ينبغي إعادة عمله.

وهناك ملاحظة أخيرة ترتبط بأهمية تقريب الأرقام العشرية (لعدد معقول) عند تقرير النتائج. وهذا لايسهل، فقط، المعادلة ولكن يجنبنا الدقة الوهمية أيضا، وربما تذكر أنه، عندما قدرنا دالة الاستهلاك في الفصل الثاني، استخدمنا بيانات بوحدات من بلايين الدولارات دون كسور، وباستخدام هذه الأرقام، قدرنا معادلة الانحدار:

$$\hat{C} = 13 + 0.89Y_d$$

ونتيجة لاستخدام الحاسوب في الوصول إلى قيم هذه المعاملات، فيمكننا الحصول على معاملات مقدرة ذات عدد عشري كبير، لذلك يمكننا أن نكتب نتائج المعادلة السابقة في الشكل:

$$\hat{C} = 13.186537 + 0.889632Y_d$$

ولكن، من الواضح أن ذلك غير مطلوب بل غير مرغوب فيه أيضا. فطالما أن بياناتنا الأساسية صحيحة، فقط، لأقرب مليون من الدولارات، فإن من غير المعقول أن نحاول التنبؤ بمستوى الاستهلاك لأقرب ألف دولار. هذا يعطينا مايطلق عليه وهم الدقة، والذي يكون غير ممكن إذا استخدمنا البيانات الخام. ينبغي علينا، عند عرض النتائج، أن نضمن تناسب عدد الأرقام العشرية المهمة الموجودة في التقرير مع مستوى الدقة التي ستسمح بها البيانات الأساسية.

(٣-٤) استخدام المتغيرات المبطة

اعتبرنا حتى الآن الشكل التالي من العلاقات:

$$C_t = a + bY_{dt} + u_t$$

$$\dot{W}_t = a + b\left(\frac{1}{R_t}\right) + u_t$$

حيث يشير الدليل السفلي t ، في تحليل السلاسل الزمنية، إلى عنصر الزمن. وتشترك جميع هذه العلاقات في خاصية واحدة، في الأقل: قيمة المتغير التابع مرتبطة بقيمة المتغير المستقل عند النقطة نفسها من الزمن (أو على مدى الفترة الزمنية نفسها). فمثلاً، افترضت نماذجنا أن الإنفاق الاستهلاكي في ١٩٥٠ م يعتمد على الدخل المتاح في ١٩٥٠ م.

ولكن، غالباً، مايتعامل الاقتصاديون مع نماذج لا تكون جميع المتغيرات فيها مرتبطة بالفترة الزمنية نفسها. افترض، مثلاً، أننا نحاول تفسير حجم الإنفاق الاستهلاكي لمجموعة من الأفراد يحصلون على دخولهم في نهاية كل شهر. قد نتوقع أن يقوم هؤلاء الأفراد بإنفاق نسبة من دخولهم خلال الشهر التالي. ولذا، تتوافر لدينا سلسلة زمنية يعتمد فيها الإنفاق في أحد الشهور على الدخل المتاح في الشهر السابق له. فإذا جعلنا t ترمز إلى الفترات بالشهور، يصبح لدينا:

$$C_t = a + bY_{d(t-1)} + u_t \quad (3.73)$$

والتي قد تشير، على سبيل المثال، إلى:

$$C_{June} = a + bY_{d(May)} + u_{June}$$

وهكذا، فإن الإنفاق الاستهلاكي t يعتمد على الدخل الذي حصل عليه خلال الفترة $(t-1)$ وتعبر عن ذلك بالقول إن الاستهلاك يتباطأ خلف الدخل بمقدار فترة زمنية واحدة، أو إن الاستهلاك يعتمد على Y_d مع فترة إبطاء واحدة.

وتوجد مجموعة أخرى من الافتراضات يمكن أن تؤدي إلى نموذج على غرار المعادلة (3.73). وتظهر هذه على النحو التالي. افترض أن:

$$C_t^p = a + bY_{dt}^e \quad (3.74)$$

حيث:

C_t^p = حجم الإنفاق الاستهلاكي المخطط للفترة المقبلة t ، وأن

Y_{dt}^e = الدخل المتوقع للفترة المقبلة t .

افترض، لتبسيط التحليل، أن:

$$Y_{dt}^e = Y_{d(t-1)} \quad (3.75)$$

يبين هذا الافتراض أن الأفراد يتوقعون أن يتماثل دخلهم في الفترة المقبلة t مع الدخل الذي حصلوا عليه في الفترة الحالية $(t-1)$. افترض، أيضا، أن:

$$C_t = C_t^p + u_t \quad (3.76)$$

حيث ترمز C_t إلى الإنفاق الاستهلاكي الفعلي في الفترة t و u_t هو الخطأ العشوائي. أي أن الإنفاق الفعلي يختلف عن الإنفاق المخطط لوجود متغير عشوائي له قيمة متوسطة صفرية. لذا ففي المتوسط، يتعادل الإنفاق الفعلي مع الإنفاق المخطط. في هذه الحال تمثل u تأثير الأحداث غير المتوقعة على حجم الإنفاق الفعلي كفاتورة زيارة طبيب مثلا. وعلى أي حال، إذا قمنا بالتعويض عن C_t^p و Y_{dt}^e في المعادلة (3.74) نحصل على:

$$C_t = a + bY_{d(t-1)} + u_t \quad (3.77)$$

والمتمثلة مع المعادلة (3.75).

وقد تفيد أنواع العلاقات المتباطئة في شرح السلوك الاستثماري والتغيرات في الأجور أيضا. فعلى سبيل المثال، لا يتخذ قرار الاستثمار في الحال، وحتى إذا اتخذ هذا القرار في الحال فإن تنفيذه يستغرق وقتا. ولهذا السبب، قد ندعي أن:

$$I_{t-1}^d = a + br_{t-1} \quad (3.78)$$

فقرار الاستثمار، I^d ، في الفترة $(t-1)$ ، يعتمد على معدل الفائدة r ، في تلك الفترة نفسها، ولكن افترض أن:

$$I_t = I_{t-1}^d + u_t \quad (3.79)$$

حيث يعتمد الإنفاق الاستثماري على قرارا الاستثمار في الفترة الزمنية السابقة (أو مع فترة إبطاء واحدة). وبالتعويض عن I_{t-1}^d في المعادلة (3.79). نحصل على علاقة تشبه العلاقة السابقة للاستهلاك أي:

$$I_t = a + br_{t-1} + u_t \quad (3.80)$$

ونترك للقارئ أن يثبت أنه، إذا دمجنا فترة إبطاء واحدة في منحني فليس حتى يصبح التغير بالنسبة المئوية في الأجور في الفترة t معتمدا على مستوى فائض الطلب على العمل في الفترة السابقة (D_{t-1}^*) ، فسوف نحصل على:

$$\dot{W}_t = a + b \left(\frac{1}{R_{t-1}} \right) + u_t \quad (3.81)$$

هل وجود مثل هذا الإبطاء يبدو منطقيا ؟

والتساؤل الذي يثور الآن هو ما إذا كان من الممكن لنموذجنا أن يعالج مشكلة تقدير العلاقة التي تتضمن متغيرات مبطأة ؟ والإجابة هي نعم. وحتى يمكننا رؤية ذلك اعتبر النموذج:

$$Y_t = a + bX_{t-1} + u_t, \quad t = 1, 2, \dots, n \quad (3.82)$$

حيث تتحقق الافتراضات الأساسية المرتبطة بالخطأ العشوائي u_t . ويتضح من هذه المعادلة أن Y تعتمد على X المتخلفة عنها بفترة إبطاء واحدة. افترض أنه يوجد لدينا عدد n من المشاهدات عن X_t ، Y_t والتي يمكن أن نعبر عنها في الجدول رقم (٥-٣).

جدول رقم (٣-٥)

Y	X
Y_1	X_1
Y_2	X_2
.	.
.	.
Y_n	X_n

لاحظ أن Y_t ليست مرتبطة بـ X_t إنها تعتمد على X_{t-1} . ولهذا السبب، ينبغي أن نضع أزواج القيم من Y و X بحيث نجعل مقابل كل قيمة من قيم Y قيمة X في الفترة السابقة لها كما في الجدول رقم (٣-٦).

جدول رقم (٣-٦)

Y	X
Y_1	X_0
Y_2	X_1
Y_3	X_2
.	.
.	.
Y_n	X_{n-1}

إذا أردنا أن نضع القيم المشاهدة لـ Y و X في شكل انتشار، فإن كل نقطة في الشكل سوف تمثل قيمة Y وقيمة X في الفترة السابقة. وهذه النقاط هي التي نوفق بها خطنا للانحدار.

لاحظ أننا، عند الانتقال من الجدول رقم (٣-٥) إلى الجدول رقم (٣-٦) خسرنا مشاهدة واحدة، حيث لن يكون بإمكاننا استخدام Y_1 طالما لا توجد لدينا مشاهدة X_0 . وبالمثل، فلن يمكننا، أيضاً، استخدام X_n طالما أننا لا نعرف Y_{n+1} . وهكذا يقلل النموذج هذا فترة الإبطاء الواحدة من حجم عيّنتنا بمقدار مشاهدة واحدة إلى $(n-1)$. أي يوجد في الجدول رقم (٣-٦) عدد $(n-1)$ ، فقط، من المشاهدات الزوجية التي يمكن استخدامها في نموذجنا.

ولتقدير علاقتنا المبطة هذه، نوجد متغيراً جديداً هو $Z_t = X_{t-1}$ (وهكذا فإن قيمة Z في أي فترة زمنية تتساوى، ببساطة، مع قيمة X في الفترة السابقة لها). ولذا يمكن إعادة كتابة نموذجنا الأساسي في المعادلة (3.82) على النحو:

$$Y_t = a + bZ_t + u_t, \quad t = 2, \dots, n. \quad (3.83)$$

وتصبح مقدراتنا لـ a, b هي:

$$\hat{b} = \frac{\sum_{t=2}^n (Z_t - \bar{Z}) Y_t}{\sum_{t=2}^n (Z_t - \bar{Z})^2} \quad (3.84)$$

$$\hat{a} = \bar{Y} - \hat{b}\bar{Z}$$

حيث:

$$\bar{Y} = \frac{\sum_{t=2}^n Y_t}{n-1}, \quad \bar{Z} = \frac{\sum_{t=2}^n Z_t}{n-1},$$

لاحظ أننا، في هذه الحسابات، نسقط كلاً من Y_1 و X_n .

مثال

لتوضيح هذه الطريقة، دعنا نعود إلى دالة الاستهلاك التي قدرناها في الفصل الثاني والتي تأخذ الشكل:

$$C_t = a + bY_{dt} + u_t$$

دعنا الآن نقدر هذه الدالة مع فترة إبطاء واحدة.

$$C_t = a + bY_{d(t-1)} + u_t \quad (3.85)$$

في المعادلة (3.85)، نفترض أن الاستهلاك في أي سنة معينة يعتمد على مستوى الدخل المتاح في السنة السابقة.

بالعودة إلى الجدول رقم (٢-٢)، ومزاوجة القيم المشاهدة للاستهلاك مع الدخل المتاح في السنة السابقة لها، نحصل على الجدول رقم (٧-٣).

جدول رقم (٧-٣)

السنة	الاستهلاك بـبلايين الدولارات	السنة	الدخل المتاح بـبلايين الدولارات
١٩٦١	٣٣٥	١٩٦٠	٣٥٠
١٩٦٢	٣٥٥	١٩٦١	٣٦٤
١٩٦٣	٣٧٥	١٩٦٢	٣٨٥
١٩٦٤	٤٠١	١٩٦٣	٤٠٥
١٩٦٥	٤٣٣	١٩٦٤	٤٣٨
١٩٦٦	٤٦٦	١٩٦٥	٤٧٣
١٩٦٧	٤٩٢	١٩٦٦	٥١٢
١٩٦٨	٥٣٧	١٩٦٧	٥٤٧
١٩٦٩	٥٧٦	١٩٦٨	٥٩٠

لاحظ أنه يتوافر لدينا الآن تسع مشاهدات، فقط، بدلا من عشر. وبتطبيق منهجنا في التقدير، نحصل على دالة الاستهلاك المتباطئة:

$$\hat{C} = \underset{(0.2)}{-20} + \underset{(44.3)}{0.98} Y_{d(t-1)}, \quad n=9, \quad R^2=0.99 \quad (3.86)$$

حيث تعبر الأعداد الموجودة بين الأقواس عن القيم المطلقة لنسب t المناظرة. من النتائج نرى أن للمعادلة (3.86) قوة تفسيرية عالية، تماما كما هو الحال بالنسبة لدالة الاستهلاك التي قدرناها في الفصل الثاني، إلا أن الحد الثابت في معادلة الاستهلاك

المبطئة لا يختلف معنويا عن الصفر (نسبة t تعادل 0.2 فقط) عند مستوى ثقة 95%. إضافة إلى ذلك، فإن الميل الحدي للاستهلاك MPC أعلى بدرجة كبيرة (0.98 مقابل 0.89). نتيجة لذلك، فإن السياسة الاقتصادية المبنية على دالة الاستهلاك المبطة قد تختلف عن تلك المؤسسة على دالة غير مبطة. من الواضح أننا نحتاج إلى طريقة تمكننا من التمييز بين هذين النموذجين. سنوجد مثل هذه الطريقة في الفصل الخامس، إضافة إلى ذلك، سوف نعالج انذاك نماذج أكثر عمومية للعلاقات المبطة.

(٣-٥) التنبؤ Prediction

نتجه في هذا الجزء إلى التطبيق الثاني الأساسي لتحليل الانحدار. في الأجزاء الأولى من هذا الفصل، أوضحنا أن نتائج الانحدار يمكن استخدامها في اختبار الفرضيات حول السلوك الاقتصادي لمختلف الوحدات الاقتصادية. كما نستخدم بدرجة الأهمية نفسها، معادلات الانحدار المقدرة في التنبؤ بتأثير أحداث معينة على المتغيرات الاقتصادية. ربما تذكر أننا ناقشنا في مقدمة الفصل الأول مشكلة المستشار الاقتصادي الذي يعمل على تقويم تأثير التخفيضات الضريبية (من أحجام مختلفة) على مستوى الإنفاق الاستثماري. فإذا افترضنا أن هذا المستشار يعلم، مثلاً، المقدار الذي يتزايد به الدخل المتاح لكل تخفيض ضريبي، فإنه يمكنه في هذه الحال استخدام العلاقة المقدرة بين C و Y في التنبؤ بتأثير مختلف التخفيضات الضريبية على حجم الإنفاق الاستهلاكي. وهنا يصبح لمعادلات الانحدار التي قدرناها دور مساعد حقيقي في تقويم الآثار الممكنة للسياسات الاقتصادية. ومن ثم، فإن تحليل الانحدار يمكن أن يساعد في فهم كمي لكيفية عمل الاقتصاد القومي، ومن ثم، التنبؤ بتأثير مختلف الاختيارات المتاحة لمصممي السياسة الاقتصادية.

والآن نرغب في فحص مشكلة التنبؤ فحوصاً أكثر دقة، افترض أن لدينا العلاقة ذات الشكل المعروف التالي:

$$Y_t = a + bX_t + u_t \quad (3.87)$$

افترض، مثلاً، أننا نعلم أصلاً أن قيمة X ستكون في فترة زمنية مستقبلية X_t . فمثلاً إذا كانت X هي مستوى الدخل المتاح، قد نفترض أن X_t هي قيمة X التي ستتج من تخفيض ضريبي معين. ومشكلتنا تتمثل في التنبؤ بقيم Y (أو Y_t) التي تناظر تلك القيمة المحددة X_t . لاحظ أن Y_t هي القيمة المستقبلية لـ Y والمناظرة لقيمة X_t المعطاة لـ X .

أول شيء ينبغي ملاحظته بالنسبة للتنبؤ هو أن Y_t ذاتها متغير عشوائي. فعلى سبيل المثال، وطبقاً لنموذجنا في المعادلة (3.87) لدينا:

$$Y_f = a + bX_f + u_f \quad (3.88)$$

حيث إن u_f هي قيمة الخطأ العشوائي في هذه الفترة المستقبلية. والآن، وفقاً لتحديداتنا المعيارية، لا يمكننا التنبؤ بـ u_f بحكم أنه متغير عشوائي ليس مرتبطاً بأي من القيم السابقة للخطأ العشوائي، أو بقيم المتغير المستقل X . وحتى إذا كنا نعلم a و b وبالتالي يمكننا حساب $(a + bX_f)$ ، فإننا لانزال غير قادرين على التنبؤ بـ Y بصورة دقيقة وذلك بسبب التأثير الذي لا يمكن التنبؤ به لـ u_f .

إضافة إلى ذلك، سيكون هناك مصدر آخر لعدم التأكد أو عدم الدقة في تنبؤنا حيث إننا، عموماً لانعرف أيًا من a و b ولذا علينا أن نستخدم تقديرات لها حساب الجزء الأول من Y_f في المعادلة (3.88) أي قيمتها المتوسطة المناظرة لـ X_f :

$$Y_f^m = a + bX_f \quad (3.88)$$

باختصار، سيكون هناك مصدران مختلفان للخطأ في تنبؤنا: الأثر الذي لا يمكن التنبؤ به للخطأ العشوائي، u_f ، واستخدام القيمة المقدرة للمعاملات a و b .

ولقد رأينا في المعادلة (3.88)، وبافتراض أن قيمة X_f معطاة، أن القيمة المستقبلية المناظرة لـ Y_f ، هي متغير عشوائي. ولذا سيكون من المرغوب فيه الحصول ليس، فقط، على تقدير النقطة والتنبؤ بـ Y_f ، ولكن بناء فترة الثقة لها أيضاً. أي أننا نرغب في الحصول على مقياس ما لدقة تنبؤنا.

سنوضح الآن طريقة لاستخدام نتائج الانحدار لتكوين التنبؤات وبناء فترات الثقة. سوف نقدم هذه الطريقة في خطوتين. وطالما أننا لانعرف أيا من a أو b ، فإننا لانعرف Y_f^m . سنتجه أولاً لاشتقاق مقدر لـ Y_f^m وآخر لتباين هذا المقدر. بعد الوصول إلى هذه المقدرات ستكون الخطوة الثانية التنبؤ بـ Y_f ذاتها. وتحديد فترات الثقة المرتبطة بها.

تقدير Y_f^m

لدينا من المعادلة (3.89) الصيغة التالية للقيمة المتوسطة لـ Y_f والمناظرة لـ X_f

وهي:

$$Y_f^m = a + bX_f \quad (3.89)$$

افترض الآن أن لدينا المقدرات \hat{a} و \hat{b} المؤسسة على عينة من Y و X ذات حجم n للفترات $t=1,2,\dots,n$ ، وطالما أن f هي فترة مستقبلية فإن $n < f$ يصبح كل من \hat{a} و \hat{b} في ظل افتراضاتنا المعتادة، مقدرات غير متحيزة. ويترتب على ذلك أنه يمكننا استخدام الصيغة $(\hat{Y}_f = \hat{a} + \hat{b}X_f)$ مقدرًا غير متحيز لـ Y_f^m ، طالما أن لدينا، وبافتراض قيمة X معطاة كمايلي:

$$\begin{aligned} E(\hat{Y}_f) &= E(\hat{a} + \hat{b}X_f) = E(\hat{a}) + [E(\hat{b})]X_f \\ &= a + bX_f = Y_f^m \end{aligned} \quad (3.90)$$

وبالعودة، مثلاً، إلى دالتنا المقدرة للاستهلاك في الفصل الثاني، افترض أن التخفيض الضريبي المقترح يرتبط بمستوى دخل متاح قدره 500 بليوناً من الدولارات. علينا حينئذ تقدير القيمة المتوقعة للإنفاق الاستهلاكي المرتبط بهذا التخفيض الضريبي على النحو:

$$\hat{C}_f^m = 13 + 0.89(500) = 13 + 445 = 458 \quad (3.91)$$

نلاحظ أن \hat{Y}_f هو تقدير النقطة للقيمة المتوسطة لـ Y_f^m ، Y_f ، والمناظرة لـ

X_f . فإذا أردنا الحصول على فترة ثقة، أو اختبار فرضيات، Y_f^m فإننا نحتاج إلى توزيع احتمالي (أو دالة) لـ \hat{Y}_f . يمكن اشتقاق هذا التوزيع، بسهولة، من نظرية أساسية في الإحصاء، تصرح بأن التوليفات الخطية من المتغيرات الطبيعية هي ذاتها موزعة توزيعاً طبيعياً. وطالما افترضنا طبيعية توزيع الأخطاء العشوائية فإن كلا من \hat{a} , \hat{b} يكون موزعاً توزيعاً طبيعياً أيضاً. وبافتراض أن قيمة X_f معطاة، فإن Y_f تصبح مؤلفاً خطياً من \hat{a} و \hat{b} ومن ثم ينبغي أن تكون متغيراً عشوائياً موزعاً توزيعاً طبيعياً أيضاً. أما القيمة المتوسطة لـ \hat{Y}_f فهي Y_f . إضافة إلى ذلك، يمكن إثبات أن تباين \hat{Y}_f هو*:

$$\sigma_{\hat{Y}_f}^2 = \sigma_u^2 \left[\frac{1}{n} + \frac{(X_f - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right] \quad (3.92)$$

حيث إن $(\bar{X} = \sum_{i=1}^n X_i/n)$ ، وإن X_1, \dots, X_n هي المشاهدات لـ X التي بنيت عليها

المقدرات \hat{a} و \hat{b} وبالتالي تكون \hat{Y}_f موزعة $N(a + bX_f, \sigma_{\hat{Y}_f}^2)$.

وقبل أن نستمر في التحليل، علينا أن نلاحظ أن تباين \hat{Y}_f يزداد مع مربع الفرق $(X_f - \bar{X})$ ، ويعني هذا إنه، كلما ابتعدت القيمة المعينة لـ X_f عن القيمة المتوسطة للعينة من المشاهدات عن \bar{X} (والتي استخدمت لبناء مقدراتنا \hat{a} و \hat{b}) يتزايد تباين مقدرنا \hat{Y}_f ، ويبدو هذا بديهياً حيث إنه، كلما ابتعدت القيمة المتنبأ بها للمتغير المستقل عن تلك القيم التي تقع ضمن مدى خبرتنا المشاهدة، انخفضت ثقتنا في دقة هذه التنبؤات. وقد نشعر بدرجة عالية من الثقة في تقدير القيمة

* لمعرفة كيفية الوصول إلى الصيغ الموجودة في هذا الجزء، انظر:

J. Johnston. *Econometric Methods*, 2nd ed. (New York: McGraw-Hill, 1972), pp. 38-43.

المتوسطة للإنفاق الاستهلاكي المرتبطة بمستوى الدخل المتاح القريب جدا من المستويات السائدة في السنوات الأخيرة. وعلى العكس من ذلك قد نشعر بدرجة كبيرة من عدم التأكد حول تقدير المستوى المتوقع من الإنفاق الاستهلاكي المرتبط بمستوى من الدخل المتاح يبلغ، بالتقريب، ضعف مستويات الدخل الحديثة.

وأخيرا، ينبغي أن نلاحظ أنه، طالما أن σ_u^2 غير معلومة عموما فإن تباين \hat{Y}_f سيكون غير معلوم، ولذا، ينبغي تقديره. ومن الواضح أن المقدّر المقترح هو:

$$\hat{\sigma}_{\hat{Y}_f}^2 = \hat{\sigma}_u^2 \left[\frac{1}{n} + \frac{(X_f - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right] \quad (3.93)$$

الذي يتبين من مناقشتنا السابقة أنه غير متحيز.

$$E(\hat{\sigma}_{\hat{Y}_f}^2) = \sigma_{\hat{Y}_f}^2 \quad (3.94)$$

وتوضح المناقشة أعلاه أنه، إذا كان σ_u^2 معلوما، فإننا نستطيع الحصول على فترات الثقة، ونختبر الفروض المرتبطة بـ Y_f^m بملاحظة أن:

$$\frac{(\hat{Y}_f - Y_f^m)}{\hat{\sigma}_{\hat{Y}_f}} \quad (3.95)$$

هو $N(0,1)$.

أما إذا كانت σ_u^2 غير معلومة، فإننا نلاحظ أن:

$$\frac{\hat{Y}_f - Y_f^m}{\hat{\sigma}_{\hat{Y}_f}} \quad (3.96)$$

هو متغير t بدرجات حرية تعادل $(n-2)$.

على سبيل المثال، إذا كانت σ_u^2 غير معلومة، تصبح فترة ثقة 95% لـ Y_f^m :

$$\left(\hat{Y}_f \pm t_{n-2;0.975} \hat{\sigma}_{\hat{Y}_f} \right) \quad (3.97)$$

وبالعودة إلى توضيحنا السابق، لدينا $\hat{C}_f = 458$ تناظر $Y_d = 500$. ولتحديد فترة ثقة 95% لـ C_f^m ، فإنه:

$$458 \pm 2.31 \left[3.4 \sqrt{\frac{1}{10} + \frac{(500 - 469)^2}{85.810}} \right] = 458 \pm 2$$

التنبؤ بـ Y_f

نتجه الآن إلى القضية ذات الأهمية الرئيسية: التنبؤ بـ Y_f ذاتها وتحديد فترات الثقة المرتبطة بها. نلاحظ أولاً، على افتراض أن قيمة X_f معطاة، أن مقدارنا (أو تنبؤنا) بالمستوى المستقبل لـ Y ، Y_f ، يكون متطابقاً مع مقدارنا Y_f^m وتحديدًا، $\hat{Y}_f = \hat{a} + bX_f$ ويتبع هذا أن المكون العشوائي الذي لا يمكن التنبؤ به في Y_f [انظر المعادلة (3.82)] له قيمة متوسطة صفرية. وبمعنى آخر فسوف نتنبأ بمستوى Y_f ببساطة عن طريق التنبؤ بقيمته المتوسطة.

في هذه الحال، يكون الخطأ في تنبؤنا هو:

$$e_f = Y_f - \hat{Y}_f \quad (3.98)$$

ويطلق على e_f خطأ التنبؤ The forecast error. وينتج عن افتراضاتنا أن خطأ التنبؤ له قيمة متوسطة صفرية:

$$\begin{aligned} E(e_f) &= E(Y_f) - E(\hat{Y}_f) \\ &= a + bX_f - a - bX_f = 0 \end{aligned} \quad (3.99)$$

افترض أن u_f ، مثل u_1 ، ...، u_n ، موزع توزيعاً طبيعياً بقيمة متوسطة صفرية. وتباين σ_u^2 . يترتب على ذلك، ومع افتراض أن قيمة X_f معطاة، أن Y_f موزع توزيعاً طبيعياً، أيضاً وأن قيمته المتوسطة هي $(a + bX_f)$ ، وأن تباينه هو σ_u^2 . وطالما أن خطأ التنبؤ e_f في المعادلة (3.98) مؤلف خطياً من المتغيرات الطبيعية (تذكر أن \hat{Y}_f

متغير طبيعي)، إذن، ينبغي أن يكون هذا الخطأ، أيضاً متغيراً طبيعياً.

والآن، وقد حددنا أن القيمة المتوسطة e_f هي الصفر، يمكننا أن نحدد تباين e_f عن طريق ملاحظة أن Y_f و \hat{Y}_f ، مع افتراض أن قيمة X_f معطاة، مستقلان. على سبيل المثال (من المعادلة 3.88) نجد أن الخطأ العشوائي الوحيد الذي يعتمد عليه Y_f هو u_f . ولكن \hat{Y}_f تعتمد على الأخطاء العشوائية u_1, \dots, u_n لأن \hat{a} و \hat{b} قد بنيت، فقط، بدلالة المشاهدات المشتركة (Y_1, X_1) ، (Y_2, X_2) ، ... و (Y_n, X_n) . وبفرض أن كلا خطأ عشوائي مستقل عن جميع الأخطاء العشوائية الأخرى فسيكون كل من Y_f و \hat{Y}_f مستقلاً، بافتراض أن قيمة X_f معطاة. ومن المعادلة (3.98) نجد أن e_f هي مولف خطي من متغيرين عشوائيين، تباينه هو (انظر ملحق الفصل الثاني):

$$\begin{aligned}\sigma_e^2 &= \sigma_{Y_f}^2 + \sigma_{\hat{Y}_f}^2 \\ &= \sigma_u^2 + \sigma_u^2 \left[\frac{1}{n} + \frac{(X_f - \bar{X})^2}{\sum (X_i - \bar{X})^2} \right] \\ &= \sigma_u^2 \left[1 + \frac{1}{n} + \frac{(X_f - \bar{X})^2}{\sum (X_i - \bar{X})^2} \right]\end{aligned}\quad (3.100)$$

وباختصار، يكون $e_f \sim N(0, \sigma_e^2)$.

وعلى نحو مشابه للتحليل السابق، نجد أن مقدراً غير متحيز لـ σ_e^2 يمكن أن يتخذ الشكل:

$$\hat{\sigma}_e^2 = \hat{\sigma}_u^2 \left[1 + \frac{1}{n} + \frac{(X_f - \bar{X})^2}{\sum (X_i - \bar{X})^2} \right] \quad (3.101)$$

ونبني فترات الثقة (والتي يطلق عليها، أحيانا، فترات التنبؤ) لـ Y_f عن طريق ملاحظة أن:

$$\frac{e_f}{\sigma_e} = \frac{Y_f - \hat{Y}_f}{\sigma_e} \quad (3.102)$$

موزعٌ توزيعاً طبيعياً معيارياً $N(0,1)$ أو، في حالة عدم معرفة σ_u^2 أن:

$$\frac{e_f}{\hat{\sigma}_e} = \frac{Y_f - \hat{Y}_f}{\hat{\sigma}_e} \quad (3.103)$$

هو متغير t بدرجات حرية قدرها $n-2$. فعلى سبيل المثال، إذا كانت σ_u^2 غير معلومة فإن فترة الثقة 95% لـ Y_f ستكون:

$$(\hat{Y}_f \pm t_{n-2;0.975} \hat{\sigma}_e) \quad (3.104)$$

من المعادلة (3.100) لاحظ أن $\sigma_e^2 > \sigma_{\hat{Y}_f}^2$ ومن المعادلتين (3.101) و (3.93)

أن $\hat{\sigma}_e^2 > \hat{\sigma}_{\hat{Y}_f}^2$ ، ويترتب على هذه النتائج أن فترة الثقة لـ Y_f ستكون أوسع من الفترة لمستوى الثقة نفسه لـ Y_f^m [قارن المعادلة (3.104) بـ المعادلة (3.97)]. وهذا هو ما ينبغي أن يكون. هناك صعوبتان في التنبؤ بـ Y_f ، الأولى هي أن المعلمتين a و b غير معلومتين، والثانية هي أن الخطأ العشوائي u_f لا يمكن التنبؤ به. أما عند التنبؤ بـ Y_f^m فتواجهنا صعوبة واحدة هي أن a و b غير معلومتين.

وعلى سبيل توضيح أخير، نعود، مرة أخرى، إلى معادلتنا المقدرة للاستهلاك، ونحسب 95% فترة ثقة لمستوى الاستهلاك. افترض أن مستوى الدخل (كما سبق) هو $Y=500$ حيث، تكون فترتنا:

$$\begin{aligned} & (\hat{a} + \hat{b}Y_d) \pm t_{n-2;0.975} \\ & = 458 \pm 2.31 \left[3.4 \sqrt{1 + \frac{1}{10} + \frac{(500 - 469)^2}{85,810}} \right] = 458 \pm 8 \end{aligned}$$

وكما لاحظنا، نجد أن فترة الثقة لـ C_f التي هي (458 ± 8) أوسع من تلك الخاصة بـ C_f^m التي هي (458 ± 2) .

(٣-٦) مثال: التقدير لمنحنى طلب

نختم معالجتنا لنموذج الانحدار البسيط بتمرين توضيحي يتضمن تقديراً لمنحنى طلب. يعرض الجدول رقم (٣-٨) بعض البيانات الفعلية حول المبيعات السنوية وأسعار الدواجن في الولايات المتحدة الأمريكية. يشير الجدول، بخاصة، إلى الاستهلاك السنوي للفرد وأسعار الاستهلاك السنوي مكمشة بالرقم القياسي لأسعار المستهلكين خلال السنوات من ١٩٤٨م إلى ١٩٦٣م. سوف نستخدم هذه البيانات لتقدير منحنى طلب على الدواجن. افترض العلاقة التالية:

$$Q_t = aP_t^b e^{u_t}, \quad (3.105)$$

حيث Q_t هي الاستهلاك الفردي بالرطل في الزمن t (خلال السنة t)، P_t هو السعر المناظر للدواجن مقاساً بالسنت Cent لكل رطل و u_t الخطأ العشوائي. على سبيل توضيح المقاييس المتضمنة، فإن قيمة $Q_t = 28.9$ ، تعني أنه، خلال السنة t كان الاستهلاك الفردي من الدجاج 28.9 رطل، وقيمة أخرى $P_t = 41.4$ تعني أن السعر المتوسط المناظر خلال تلك السنة كان 41.4 سنتاً لكل رطل. يشير الجدول رقم (٣-٨) إلى أن هذه الأرقام تناظر عام ١٩٥٩م.

افترض الآن أن الخطأ العشوائي u_t يحقق الافتراضات كافة لنموذجنا المعتاد، حيثئذ وكما أشرنا من قبل فيما يتصل بالمعادلة (3.43)، يمكن تفسير المعلمة b في النموذج [كما في المعادلة (3.105)] بأنها مرونة. وفي هذه الحال، تصف b النسبة المثوية المتوقعة للتغير في الاستهلاك الفردي السنوي من الدجاج المصاحب لمعدل تغير في السعر قدره ١٪.

وكما رأينا من قبل في هذا الفصل، يمكننا استخدام التحويلة اللوغاريتمية لجعل النموذج (3.105) يأخذ الشكل الخطي:

$$\ln(Q_t) = A + b \ln(P_t) + u_t \quad (3.106)$$

حيث $A = \ln(a)$ ، وأخذت اللوغاريتمات للأساس e . وهذا يقترح، كما نوقش من قبل، أننا نأخذ، ببساطة، اللوغاريتم الطبيعي لكل من المتغيرين ثم نجري انحدارا للوغاريتم المتغير التابع $[\ln(Q_t)]$ على لوغاريتم المتغير المستقل $[\ln(P_t)]$ ، لكي نحصل على تقديرات لـ A و b . (مثلا، \hat{A} و \hat{b}). حيثئذ، يكون تقديرنا (المتحيز) لـ a هو \hat{A} .

جدول رقم (٣-٨) الاستهلاك الفردي والسعر المكش للدجاج (1948-1963)

السنة	الاستهلاك (ببلايين الدولارات)	السعر المكش لرطل (بالسنت)
١٩٤٨	١٨,٣	٧٥,٤
١٩٤٩	١٩,٦	٧٥,٤
١٩٥٠	٢٠,٦	٧١,٨
١٩٥١	٢١,٧	٦٨,٠
١٩٥٢	٢٢,١	٦٦,٠
١٩٥٣	٢١,٩	٦٥,٠
١٩٥٤	٢٢,٨	٥٦,٤
١٩٥٥	٢١,٣	٥٨,٧
١٩٥٦	٢٤,٤	٥٠,٤
١٩٥٧	٢٥,٥	٤٧,٦
١٩٥٨	٢٨,٢	٤٥,٨
١٩٥٩	٢٨,٩	٤١,٤
١٩٦٠	٢٨,٢	٤١,٤
١٩٦١	٣٠,٣	٣٧,٠
١٩٦٢	٣٠,٢	٣٨,٦
١٩٦٣	٣٠,٦	٣٧,٦

كُمش السعر باستخدام الرقم القياسي لأسعار المستهلكين ١٩٥٧ - ١٩٥٩ = ١٠٠

المصدر: Frederick V. Waugh. *Demand and Price Analysis-Some Examples from Agriculture* (Washington, D.C : U.S. Department of Agriculture, Technical Bulletin 1316, Nov. 1964), Table 5-1, p. 39

وبتطبيق هذا المنهج على المعادلة (3.106) على البيانات الموجودة في الجدول (٣-٨) نحصل على:

$$\ln(Q_i) = 5.87 - 0.68 \ln(P_i) = \quad (3.107)$$

$(01.3) \quad (0.03)$

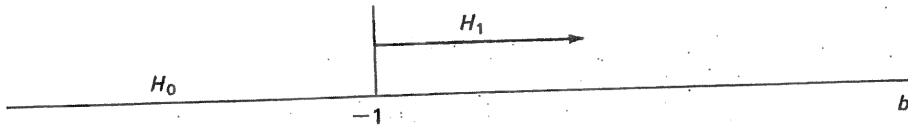
$$R^2 = 0.97$$

حيث الأرقام داخل الأقواس تحت المعاملات المقدرة هي الأخطاء المعيارية المقدرة المناظرة. وبهذا، يكون تقدير مرونة النقطة للطلب على الدواجن -0.68 وتقديرنا للمعلمة a بدرجة دقة لثلاثة أرقام هو $(e^{5.57} = 354)$.

والآن، نوضح بعض الطرق التي عرضناها خلال هذا الفصل بدلالة نتائج المعادلة (3.107). افترض، على سبيل المثال، أننا نهتم باختبار الفرضية بأن الطلب على الدجاج غير مرتبط بسعره عند مستوى معنوية 5% مقابل الفرضية البديلة بأن الطلب على الدجاج يتأثر بالسعر. حينئذ تكون فرضية العدم $H_0: b=0$ والفرضية البديلة $H_1: b \neq 0$ ، وبأخذ القيمة المطلقة لتقدير b في المعادلة (3.107) وبقسمته على الخطأ المعياري المقدّر المناظر $(0.68/0.03)$ ، ينتج عنه نسبة عالية جداً لـ t تزيد على 20. وباستخدام قاعدتنا التجريبية للحساب يمكننا أن نستنتج مباشرة أن الفرضية $H_0: b=0$ سترفض وفقاً للنتائج في المعادلة (3.107).

وهناك توضيح آخر، افترض أننا نهتم باختبار فرضية أن الطلب على الدجاج مرّن سعرياً عند واحد في المائة مستوى معنوية، مقابل الفرضية البديلة بأنه غير مرّن. من مبادئ الاقتصاد الجزئي، نعرف أن الطلب يكون مرّنًا سعرياً إذا كان معامل المرونة أقل من (-1) أو أكبر من واحد (بالقيمة المطلقة). وبدلالة نموذج الطلب بالمعادلة (3.105)، يناظر هذا أي قيمة لـ b (بحيث تكون $b < -1$). وهكذا فإن فرضيتنا للعدم والفرضية المقابلة في هذه الحالة سيكونان $H_0: b < -1$ و $H_1: b > -1$. وبخلاف التوضيحات المعطاة في الأجزاء السابقة لهذا الفصل، فإن فرضيتنا للعدم لا تحدد قيمة معينة لـ b . وعلى الرغم من ذلك، فإن هذه الفرضية يمكن اختبارها بوساطة منهج فترة الثقة. ولنرى ذلك، لاحظ أن H_1 هي اختبار الذيل

الواحد، الذي يفيد أن قيمة b أكبر من جميع القيم الممكنة لها والمحددة بوساطة H_0 . ويصور الشكل (٨-٣) هذه الحالة. وانطلاقاً من نقاشنا السابق في هذا الفصل نبنى فترة ثقة بذيل واحد ويحد أدنى، مثلاً LB . وستأخذ هذه الفترة شكل $b \geq LB$. إذا كانت $LB \geq -1$ فإننا نرفض H_0 ، وإذا كانت $LB < -1$ فإننا نقبل H_0 . لاحظ أن منهج الاختبار هذا يؤدي آلياً إلى قبول H_0 ، وما رأيانه هو نشوء صعوبة إذا كانت فرضيتنا للعدم تأخذ الشكل $(H_0 : b \neq 0)$.



شكل (٨-٣)

وطريقة التحليل هذه مباشرة وواضحة. فباستخدام الجدول الإحصائي رقم ٢ لتوزيع t بدرجات حرية قدرها $(16-2=14)$ ، نجد أن فترة الثقة 99% ذات الذيل الواحد هي:

$$b > (\hat{b} - t_{14;0.99} \hat{\sigma}_{\hat{b}}) = -0.68 - (2.624)(0.03) = -0.76$$

وطالما أن $-0.76 > -1$ ، فإننا نرفض H_0 ونقبل الفرضية البديلة بأن الطلب على الدجاج ليس مرناً سعرياً.

اعتبر مرة أخرى النتائج في المعادلة (3.107)، ولاحظ أن المعادلة المقدرة «توفق» البيانات الملاحظة دائماً. فالمعادلة تفسر 97% من التغير المشاهد في الاستهلاك الفردي السنوي المتوسط من الدجاج. ولكنك قد لاتشعر بالارتياح، وينبغي لك، في الأقل، لسبيين، الأول: تفسر معادلتنا التغيرات السنوية في استهلاك الدجاج الناجمة عن تغيرات السعر، ولكننا نعلم أنه على مدى الفترة من ١٩٤٨ إلى ١٩٦٣م زادت دخول المستهلكين، فإذا كان الدجاج سلعة عادية، فإننا نتوقع أن

ارتفاع الدخل ينبغي أن ينسب إليه بعض الزيادة في استهلاك الدجاج خلال تلك الفترة. ولكن المعادلة (3.105) تهمل أي إشارة إلى تأثير ارتفاع الدخل. وفي هذه الحال، فإن حذف متغير مهم (الدخل) ينسب إلى تأثير السعر ليس، فقط، على استهلاك الدجاج ولكن، أيضا، على الدخل. لهذا السبب فإن قيمتنا المقدرة تبالغ في تأثير متغير السعر على المشتريات الاستهلاكية من الدجاج. والثاني: أننا، في منهجنا للتقدير، لم نأخذ في الحسبان جانب العرض من السوق، حيث إن الأسعار المشاهدة والكميات ليس نتاجا لتأثير الطلب، فقط، ولكنها أيضا، نتاج تفاعل الطلب والعرض. ينبغي أن ندخل ذلك صراحة في نموذج الانحدار، وهذا ماستناوله في الفصول المتبقية.

أسئلة

- ١- يُدّعي الآن أن الأداء في اختبارات I.Q قد تحسن في السنوات الأخيرة، وأن الوسط الحسابي للأداء الآن فوق ١٠٠. افترض أن الدرجات المتحصل عليها في I.Q موزعة توزيعا طبيعيا عبر المجتمع الإحصائي. فإذا كانت لدينا عينة مكونة من ١٠٠ اختبار حيث القيمة المتوسطة ١١٠ والتباين المقدّر هو ٤. اختبر فرضية أن القيمة المتوسطة للمجتمع الإحصائي I.Q أكبر من ١٠٠ عند مستوى معنوية ٥٪.
- ٢- اشرح لماذا يكون الوسط الحسابي للمجتمع المحسوب على أساس القيمة المتوسطة لعينة عشوائية مكونة من ٣٠ مشاهدة أفضل من تلك العينة المكونة من ٢٠ مشاهدة. هل كلا المقدرين غير متحيزين.
- ٣- افترض أن أسبوع العمل المعتاد في الصناعة هو ٤٠ ساعة. وفرضيتنا هي أنه، إذا اختلفت الساعات عن ٤٠ فسوف تميل للعودة إلى ٤٠ مرة أخرى. احدى طرق تكوين ذلك هو $\Delta H_t = B + \alpha(40 - H_{t-1}) + u_t$ حيث $\Delta H_t = H_t - H_{t-1}$ ، وهكذا فإن $H_t = (B + 40\alpha) + (1 - \alpha)H_{t-1} + u_t$. افترض أننا نقدر الانحدار التالي من البيانات ربع السنوية للولايات المتحدة الأمريكية:

$$\hat{H}_t = 5 + 0.875 H_{t-1}, \quad R^2 = 0.98$$

(0.7) (0.15)

حيث تظهر الانحرافات المعيارية المقدرة داخل الأقواس تحت المعاملات المقدرة. هل ينبغي أن نقبل أو نرفض الفرضية بأن التغير في الساعات يعتمد على الانحراف عن ٤٠؟ لماذا؟

٤- وضع السيد (أ) نظرية تقول إن الوسط الحسابي لأطوال الأفراد هو ٧٠ بوصة بينما يجادل السيد (ب) بأن نظرية السيد (أ) تبالغ في تأثير عوامل معينة، ومن ثم، فهي تبالغ في القيمة المتوسطة. افترض أن الأرقام التالية هي نتائج عينة عشوائية ذات حجم ٤ (٦٢ ، ٧٢ ، ٧٤ ، ٦٤). افترض أن دالة الكثافة الاحتمالية موضع الاعتبار طبيعية وذات $\sigma^2 = 4$. اختبر نظرية السيد (أ) عند مستوى معنوية ٥٪.

٥- ما أهمية افتراضنا بأن الخطأ العشوائي موزع توزيعاً طبيعياً؟

٦- افترض أن الوسط الحسابي لأطوال الأفراد في الجانب الشرقي من الولايات المتحدة الأمريكية هو ٦٧ بوصة. افترض أن صانعا للملابس الجاهزة في الشرق يريد أن يفتح متجرًا لبيع الملابس في الجانب الغربي من الولايات المتحدة الأمريكية، وهو يعتقد أن الأفراد في الغرب أطول من الأفراد في الشرق. وإذا كان الأمر كذلك فإن عليه أن يصنع ملابس جاهزة أطول بعض الشيء. افترض أن هذا يتطلب إعادة تجهيز مكلف للآلات. وافترض أيضاً أن هذه الفرضية المرتبطة بالطول النسبي قد اختيرت بدلالة عينة عشوائية من أفراد إختيروا من الغرب. ناقش نتائج الخطأ من النوع الأول والخطأ من النوع الثاني.

٧- إشرح لماذا لا يكون نموذج الانحدار الخطي مقيداً تقييداً كبيراً، على الرغم من أن كثيراً من العلاقات الاقتصادية غير خطية. اختر تحويلًا ملائماً ثم اشتق مصفوفة المشاهدات للنموذج:

$$Y_t = a + b \left(\frac{1}{1 - X_t} \right) + u_t$$

عندما تكون $n=3$ ، ومشاهداتنا عن X و Y هي $Y_3 = 12$ ، $Y_2 = 10$ ، $Y_1 = 1$ و $X_1=0$ ، $X_2=0.1$ و $X_3=0.5$.

٨- افترض أن أحد الأفراد قدر معادلة للاستهلاك وظهرت النتائج على النحو:

$$\hat{C} = \underset{(3.1)}{15} + \underset{(18.7)}{0.81} Y_d \quad n=19 \quad R^2=0.99$$

حيث الأرقام الموجودة داخل الأقواس هي نسب t .

(أ) استخدم نسبة t لاختبار الفرضية بأن Y_d هو متغير مهم معنوياً.

(ب) حدد الانحرافات المعيارية المقدرة لمقدرات المعلمات.

(ج) كون فترة ثقة 95% لمعامل Y_d . هل تشتمل هذه الفترة على الصفر؟

٩- اعتبر الموقف الذي تؤدي فيه الزيادة في معدل المنافع للضمان الاجتماعي

(مقاسة بالمدفوعات لكل أسرة في الشهر) إلى زيادة في الطلب على الضمان

الاجتماعي عن طريق إحلال الأفراد للفراغ محل العمل. افترض، أيضاً، أن

الزيادة في الطلب على الضمان الاجتماعي تؤدي، بدورها، من خلال الضغوط

السياسية، للزيادة في معدل المنافع للضمان الاجتماعي في الفترة التالية. عبر

عن هذه العلاقات باستخدام نموذج مكون من معادلتين.

١٠- اعتبر الحال التي يعتمد فيها عدد المنشآت الاقتصادية التي تستوطن ولاية معينة

على معدل الضريبة النسبي في هذه الولاية. افترض، أيضاً، أنه، على الرغم

من وجود منافع ضريبية للمقيمين في الولاية، فإن زيادة عدد المنشآت التي

تستوطن ولاية معينة تؤدي إلى زيادة معدل التلوث بها. عبر عن العلاقات بين

مكان توطن المنشأة والتلوث بدلالة نماذج الانحدار.

١١- اعتبر نموذج الانحدار المعتاد التالي:

$$Y_t = a + bX_t + u_t$$

افترض أنه لا يمكننا قياس X_t ، وافترض، بدلاً من ذلك أننا نشاهد المتغير Z_t

حيث $Z_t = 5 - 3X_t$ ، ما معلمات النموذج التي تربط بين Y_t و Z_t ؟

تحليل الانحدار المتعدد

استكشفنا في الفصول السابقة العلاقة الإحصائية بين متغيرين، ففي حالة الإنفاق الاستهلاكي (مثلا) استخدمنا النموذج:

$$C_t = a + bY_{dt} + u_t, \quad (4.1)$$

وطورنا الطرق التي يمكننا عن طريقها تقدير كل من a و b ، واختبار الفرضيات حول هذه العلاقة. إلا أننا، عادة، نجد أن العلاقات الاقتصادية أكثر تعقيدا من هذا النموذج (4.1) بمعنى أن قيمة متغير معين كالاستهلاك تعتمد ليس، فقط، على متغير واحد، وإنما على مجموعة كاملة من المتغيرات المستقلة.

افترض، على سبيل المثال، أن مستوى الإنفاق الاستهلاكي في الفترة t يعتمد ليس فقط، على مستوى الدخل المتاح الحالي، ولكن، أيضا، على قيمة الأصول السائلة $(A_t)^*$ وعلى الدخل المتاح في الفترة السابقة $Y_{d(t-1)}$. فإذا كان ما يمتلكه المستهلك من الأصول السائلة كبير جدا غير عادي (كما كان الحال عندما شارفت الحرب العالمية الثانية على الانتهاء)، فسوف نتوقع أن يكون الإنفاق الاستهلاكي أكبر، نوعا ما، عن ذلك المرتبط عادة بمستوى الدخل السائد. وعلى العكس إذا كان المخزون من الأصول السائلة منخفضا انخفاضا غير عادي، فإن

* نعني بالأصول السائلة ما يملكه الفرد من النقود والودائع الآجلة والمدخرات، وانصبة القروض، والسندات الحكومية.

المستهلكين قد يقللون إنفاقهم الاستهلاكي بعض الشيء من أجل استكمال النقص في ممتلكاتهم من الأصول السائلة. وقد يؤدي الدخل السابق دوراً في تحديد المستويات الحالية للإنفاق الاستهلاكي، فمثلاً قد تكون بعض النفقات التي تتم في الفترة الحالية محفزة بوساطة المستويات السابقة للمعيشة. كما قد ترتبط بمستوى الدخل في الفترة السابقة مباشرة. فعلى سبيل المثال، مع بقاء العوامل الأخرى على حالها، إذا كان الدخل أعلى في الفترة السابقة مباشرة، فإن من المرجح أن يرتفع مستوى الاستهلاك الحالي. وكل هذا يعني أنه لدينا الآن دالة استهلاك معقدة مثل:

$$C_t = a + b_1 Y_{dt} + b_2 A_t + b_3 Y_{d(t-1)} + u_t. \quad (4.2)$$

ومشكلتنا الآن هي تقدير كيفية اعتماد الاستهلاك على جميع هذه المتغيرات المستقلة، وبمعنى آخر، ينبغي أن توجد طريقة يمكن بوساطتها تقدير قيم a ، b_1 ، b_2 و b_3 .

ومرة أخرى، نود معرفة تباين مقدراتنا حتى نستطيع الحصول على مقياس لدقة هذه المقدرات. وعلى سبيل تعميم مباشر لما عرفناه من قبل، سنفترض أن لدينا مجموعة من القيم المشاهدة لمتغيرات نموذجنا، فعلى سبيل المثال، وبالمقابل مع المعادلة (4.2) فإن لدينا معلومات تقابلها في الجدول رقم (٤-١).

للوهلة الأولى، تظهر طبيعة مشكلتنا مختلفة بعض الشيء وأكثر تعقيداً مقارنة بحالة الانحدار البسيط. وبالتحديد، ففي حالة الانحدار البسيط، لدينا قيم مشاهدة عن المتغير التابع ومتغير مستقل واحد، وكانت مشكلتنا، ببساطة، هي تقدير كيف يتغير الأول مع المتغير الثاني، والآن، لدينا مشاهدات عن مجموعة كاملة من المتغيرات المستقلة، ونواجه مهمة أكثر صعوبة تتمثل في معرفة آثار المتغيرات المستقلة كافة. أي أنه لتحديد، تأثير كل متغير مستقل، ينبغي علينا أن نزل، بدرجة ما، تأثيره على المتغير التابع عن تأثير باقي المتغيرات المستقلة الأخرى.

جدول رقم (١، ٤) القيم المشاهدة (بلايين الدولارات)

السنة	الاستهلاك	الدخل المتاح	الأصول المالية	الدخل المتاح من السنة السابقة
	(C _t)	(Y _{dt})	(A)	Y _d (t-1)
١٩٦٠	٥٢٥,٢	٣٥٠,٠	٣٩٩,٢	٣٣٧,٣
١٩٦١	٣٣٥,٢	٣٦٤,٤	٤٢٤,٦	٣٥٠,٠
١٩٦٢	٣٥٥,١	٣٨٥,٣	٤٩٩,٠	٣٦٤,٤
١٩٦٣	٣٧٥,٠	٤٠٤,٦	٤٩٥,٤	٣٨٥,٣
١٩٦٤	٤٠١,٢	٤٣٨,١	٥٣٠,٥	٤٠٤,٦
١٩٦٥	٤٣٢,٨	٤٧٣,٢	٥٧٣,١	٤٣٨,١
١٩٦٦	٤٦٦,٣	٥١١,٩	٦٠١,٥	٤٧٣,٢
١٩٦٧	٤٩٢,١	٥٤٦,٣	٦٥٠,٤	٥١١,٩
١٩٦٨	٥٣٥,٨	٥٩١,٢	٧٠٩,٦	٥٤٦,٣
١٩٦٩	٥٧٧,٥	٦٣١,٦	٧٣١,٦	٥٩١,٢

المصدر: التقرير الاقتصادي للرئيس (واشنطن، مكتب الطباعة الحكومية في الولايات المتحدة الأمريكية،

فبراير ١٩٧١م) الصفحات ١٩٧، ٢٠٤ و ٢٦٢.

وعلى الرغم من هذا التعقيد الظاهري، فإننا نؤكد في البداية أن تحليل الانحدار المتعدد (أي الحال التي يكون لدينا فيها أكثر من متغير مستقل واحد) هو تعميم مباشر للتحليل الثنائي المتغيرات. سنعرض في البداية نموذج انحدار متعدد يحقق الافتراضات المرتبطة بخصائص الخطأ العشوائي كافة التي كونها في حال انحدار المتغيرين. بعد ذلك، سوف نتبع طريقة المتغير المساعد التي، عن طريقها نضع افتراضاتنا بوصفها شروطاً على مقدار الخطأ العشوائي. سيستج عن ذلك مجموعة من المعادلات الطبيعية، وبحل هذه المعادلات، نحصل على مقدرات غير متحيزة لكل المعاملات في نموذجنا. وبهذا، يكون منهجنا تقريباً هو المنهج نفسه المستخدم في الفصول السابقة. فإذا كنت قد فهمت حالة الانحدار البسيط فلن تواجه صعوبة كبيرة في فهم تحليل الانحدار المتعدد.

(٤-١) نموذج الانحدار المتعدد

عموماً، دعنا نعتبر نموذج الانحدار التالي:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + \dots + b_k X_{kt} + u_t, \quad t=1, \dots, n, \quad (4.3)$$

حيث:

$$Y_t = \text{المشاهدة رقم } t \text{ عن المتغير التابع.}$$

$$X_{it} = \text{المشاهدة رقم } t \text{ عن المتغير المستقل } X_i, \text{ حيث } (i=1,2,\dots,k) \text{ إذا كان لدينا}$$

مشاهدات مستقلة عددها k .

$$u_t = \text{القيمة } t \text{ للخطأ العشوائي، وأخيراً.}$$

$$b_i = \text{معامل المتغير المستقل رقم } i.$$

سوف يكون من الملائم بدءاً من الآن استخدام b_0 (بدلاً من a) لرمز إلى الحد الثابت في معادلة الانحدار المتعدد حتى تأخذ الملمات، كافة في المعادلة (4.3) الشكل b 's. سنعرض الآن قائمة بالافتراضات التي نكونها لنموذج الانحدار المتعدد. وتتطلب غالبية هذه الافتراضات تعليقا محدودا طالما أنها الافتراضات نفسها التي فرضناها في حالة الانحدار البسيط:

١ - القيمة المتوقعة أو المتوسطة للخطأ العشوائي هي الصفر

$$E(u_t) = 0.$$

٢ - تباين الخطأ العشوائي ثابت، ولذا يكون مستقلاً عن

$$E(u_t - 0)^2 = E(u_t)^2 = \sigma_u^2.$$

٣ - قيم الخطأ العشوائي مستقلة عن بعضها بعضاً، لذا، يكون التباين بين الأخطاء

العشوائية المناظرة لأي مشاهدين (u_t, u_s) صفراً،

$$\text{cov}(u_t, u_s) = E(u_t u_s) = 0.$$

٤ - استقلال الخطأ العشوائي عن جميع قيم المتغيرات المستقلة. وبالتحديد نفترض

أن u_t مستقل عن X_{1s}, \dots, X_{ks} لجميع s, t . ويتبع عن ذلك أن التباين بين

الخطأ العشوائي u_t وكل واحد من المتغيرات المستقلة لمعادلتنا للانحدار (4.3)

هو الصفر. ومرة ثانية، فإن هذا يعني أنه، سواء وضعت قيم المتغيرات

المستقلة بوساطة «الباحث» الذي يجري التجارب أو بوساطة «الاقتصاد» فإن تلك القيم لا تتأثر بأي طريقة كانت على قيم الخطأ العشوائي. واصطلاحاً يمكن التعبير عن شرط التباين المشترك بين u_t وكل من X_{it} على النحو:

$$\text{cov}(u_t, X_{it}) = E[u_t(X_{it} - \mu_{Xi})] = E(u_t X_{it}) - \mu_{Xi} E(u_t) = E(u_t X_{it}) = 0$$

حيث $E(X_{it}) = \mu_{Xi}$

٥ - لا يوجد ارتباط خطي تام بين المتغيرات المستقلة. بمعنى أن أي المتغيرات المستقلة ليس توليفة خطية من المتغيرات الأخرى. على سبيل المثال، تستبعد هذه القاعدة علاقات مثل:

$$\begin{aligned} \text{a. } X_{1t} &= 3 - 2X_{2t} + 17X_{3t}; \\ \text{b. } X_{4t} &= (X_{1t} + X_{2t} + X_{3t}) / 3; \text{ or} \\ \text{c. } X_{2t} &= 3X_{8t}. \end{aligned}$$

ولكننا لانستبعد العلاقات غير الخطية، فمثلاً إذا كانت $X_{1t} = Y_{2t}^2$ ، وإذا كانت $X_{3t} = X_{5t} X_{6t}$ ، فإن افتراضنا سيظل صحيحاً.

تعرفت من قبل على الافتراضات الأربعة الأولى من هذه الافتراضات من نموذج الانحدار البسيط، وتبريرها هو التبرير نفسه الذي اوردناه في الحالة السابقة. فإذا لم تكن متأكداً من سبب تكوين هذه الافتراضات الأربعة فعليك أن تجد ذاكرك بالعودة إلى المناقشة الموجودة في حالة الانحدار البسيط في الفصل الثاني. أما الافتراض الجديد الذي أضفناه هنا فهو الافتراض الخامس. وهو، في الحقيقة، امتداد للافتراض الذي كونه في الفصل الثاني بأن المتغير المستقل X_t في حالة انحدار المتغيرين ينبغي أن يكون له، في الأقل، قيمتان مختلفتان. رأينا في الفصل الثاني في حال انحدار المتغيرين إذا كانت قيمة X لا تتغير، أي أن $X_t \equiv X_0$ ، فإننا فقط نستطيع تقدير معلمة واحدة، والتي نطلق عليها A ، تعتمد على كل من الحد الثابت الأصلي، a ، وعلى ماهو، في الحقيقة، حد ثابت أوجد بوساطة

القيمة غير المتغيرة لـ X_{it} و bX_0 ، أي $A = a + bX_0$. وباختصار، إذا كانت قيمة المتغير المستقل لا تتغير مطلقاً، فإن تأثيره على Y لا يمكن فصله عن تأثير الحد الثابت الأصلي.

في حال الانحدار المتعدد هذه، نرغب في ضمان ليس، فقط، أن b_i أو أن تأثير X_{it} على Y ، يمكن عزله عن الحد الثابت b_0 ، ولكن أيضاً يمكن عزله عن تأثير المتغيرات المستقلة الأخرى كافة. وتظهر قوة الافتراض الخامس في أنه يضمن مثل هذا العزل. وسنبين هذا بدقة في المبحث (٤-٢)، ولكننا نرغب الآن في توضيح لماذا يمكن أن يترتب على انتهاك الافتراض الخامس.

افترض أن X_{1t} تأخذ قيمة متغيرة ولكنها تتعادل دائماً مع X_{2t} . حيثُذ (ومع بقاء العوامل الأخرى على حالها)، إذا تزايد X_{1t} بمقدار وحدة واحدة فإن Y_t سوف يتغير بمقدار $(b_1 + b_2)$ وحدة لأنه إذا كان X_{it} يتعادل دائماً مع X_{2t} ، فإن X_{2t} يتزايد أيضاً، بمقدار وحدة واحدة! يوحي هذا بأن التأثير المؤلف لـ X_{1t} و X_{2t} (وهو $b_1 + b_2$) هو الذي يمكن تقديره إذا كانت قيم $X_{1t} = X_{2t}$. وببساطة لا توجد طريقة يمكن بواسطتها عزل تأثير X_1 من تأثير X_2 على Y . ويتبع هذا بسبب أنه، إذا كانت $X_{1t} = X_{2t}$ ، فإن نموذجنا الأساسي بالمعادلة (4.3) يمكن أن تعاد كتابته على النحو:

$$Y_t = b_0 + BX_{1t} + b_3X_{3t} + \dots + b_kX_{kt} + u_t, \quad (4.4)$$

حيث إن $B = (b_1 + b_2)$. إلا أنه، إذا كان المتغيران المستقلان متساويين دائماً مع بعضهما بعضاً فإنه يمكن حذف واحد من المتغيرين من نموذجنا دون خسارة في المعلومات. وحيثُذ، يمكن اعتماد النموذج الناتج (أو المختزل) والذي يحتوي على $(k-1)$ متغيرات مستقلة، وسيحتوي هذا النموذج على المعلمة B السابق الإشارة إليها، التي تعبر عن الأثر المؤلف للمتغيرين المستقلين الأصليين في المعادلة. لاحظ أن المعادلة (4.4) لا تقترح عدم قابلية b_0 ، b_3 ، \dots ، b_k للتقدير ومرة أخرى سنعود إلى هذه المشكلة بدقة أكثر في المبحث (٤-٢).

وهكذا، فإن نموذج الانحدار المتعدد مشابه جدا لنموذج الانحدار البسيط، فهو يصف العلاقة الدالية الخطية والتي تتحدد عن طريقها قيمة المتغير التابع بوساطة قيم مجموعة من المتغيرات المستقلة. والآن ينبغي علينا أن نوجد طريقة لتقدير قيم المعلمات في هذه العلاقة.

(٤-٢) التقدير بوساطة المتغيرات المساعدة

تذكر أن طريقة المتغيرات المساعدة تتضمن فرض مجموعة من الشروط على مقدر الخطأ العشوائي التي تقترحها افتراضات نموذج الانحدار. ففي حالة الانحدار البسيط فرضنا الشرطين التاليين:

الافتراض المناظر الشرط

$$1. \sum \frac{\hat{u}_i}{n} = 0, \text{ أو } \sum \hat{u}_i = 0 \quad E(u_i) = 0$$

$$2. \sum \frac{(X_i \hat{u}_i)}{n} = 0, \text{ أو } \sum (X_i \hat{u}_i) = 0 \quad E(X_i u_i) = 0$$

وبمساعدة هذين الشرطين، أوجدنا معادلتين طبيعيتين، حللناهما للحصول على قيم \hat{a} و \hat{b} . وهنا سوف نستخدم المنهج نفسه ولكن قبل أن نقوم بذلك، ينبغي علينا أن نستخدم بعض تعريفاتنا الأساسية السابقة ضمن إطار الانحدار المتعدد. فعلى سبيل المثال، يعتمد منهجنا في التقدير اعتمادا حاسما على مقدر الخطأ العشوائي \hat{u}_i . لذلك ينبغي علينا تعريف \hat{u}_i ضمن إطار نموذج الانحدار المتعدد. إذا كان نموذجنا للانحدار هو (4.3)، فإن القيمة المتوسطة لـ Y_i المرتبطة بقيم

X_{1i}, \dots, X_{ki} هي:

$$Y_i^m = b_0 + b_1 X_{1i} + \dots + b_k X_{ki}. \quad (4.5)$$

وكما هو الحال في حالة الانحدار البسيط يمكن كتابة Y بوصفها مجموع كل من قيمتها المتوسطة والخطأ العشوائي:

$$Y_t = Y_t^m + u_t. \quad (4.6)$$

وإذا عرفنا b_0, b_1, \dots, b_k ، يمكننا أن نشق قيمة u_t على النحو:

$$u_t = Y_t - Y_t^m. \quad (4.7)$$

لاحظ أن المعادلتين (4.6) و (4.7) تتماثلان مع نظيرتهما في حال الانحدار البسيط. افترض أن لدينا مقدرات لـ b_0, b_1, \dots, b_k مثلاً $\hat{b}_0, \hat{b}_1, \dots, \hat{b}_k$ في ضوء المعادلة (4.5)، يصبح مقدرنا للقيمة المتوسطة لـ Y_t :

$$\hat{Y}_t = \hat{b}_0 + \hat{b}_1 X_{1t} + \dots + \hat{b}_k X_{kt}, \quad (4.8)$$

لقد أسقطنا، أيضاً في هذه المرة الرمز العلوي m من أجل التبسيط. وحيث أن يكون مقدرنا للخطأ العشوائي المقترح بوساطة المعادلة (4.7) هو:

$$\hat{u}_t = Y_t - \hat{Y}_t \quad (4.9)$$

$$= Y_t - \hat{b}_0 - \hat{b}_1 X_{1t} - \dots - \hat{b}_k X_{kt}.$$

يمكننا إعادة كتابة المعادلة (4.9) على النحو:

$$Y_t = \hat{Y}_t + \hat{u}_t, \quad (4.10)$$

أو بطريقة متكاملة على النحو:

$$Y_t = \hat{b}_0 + \hat{b}_1 X_{1t} + \dots + \hat{b}_k X_{kt} + \hat{u}_t. \quad (4.11)$$

باختصار، تناظر تعريفاتنا لكل من \hat{Y}_t و \hat{u}_t بدقة المقيدين أنفسهما في حالة نموذج الانحدار البسيط في الفصل الثاني. والآن، طالما يتوافر لنا تعريفات كافية فستجده إلى التقدير.

المعادلات الطبيعية

لإيجاد المعادلات الطبيعية في حالة نموذج الانحدار المتعدد، ستتبع المنهج نفسه الذي اتبعناه في إيجاد تلك المعادلات في حالة نموذج انحدار المتغيرين. والآن، وبافتراض وجود عدد k من المتغيرات المستقلة، سيصبح لدينا $(k+1)$ شروط نفرضها على مقدر الخطأ العشوائي وبالتحديد، لدينا:

$$1. \sum \frac{\hat{u}_t}{n} = 0, \text{ أو } \sum \hat{u}_t = 0 \quad E(u_t) = 0$$

$$2. \sum \frac{(X_{1t} \hat{u}_t)}{n} = 0, \text{ أو } \sum (X_{1t} \hat{u}_t) = 0 \quad E(X_{1t} u_t) = 0$$

$$3. \sum \frac{(X_{2t} \hat{u}_t)}{n} = 0, \text{ أو } \sum (X_{2t} \hat{u}_t) = 0 \quad E(X_{2t} u_t) = 0$$

$$k+1. \sum \frac{(X_{kt} \hat{u}_t)}{n} = 0, \text{ أو } \sum (X_{kt} \hat{u}_t) = 0 \quad E(X_{kt} u_t) = 0$$

لاحظ أن لدينا الآن (k+1) شروط و (k+1) معلمات هي b_0, \dots, b_1, b_k التي تظهر في نموذجنا للانحدار.

دعنا الآن نفحص كيف تولد هذه الشروط عدد (k+1) من المعادلات الطبيعية.

إذا جمعنا أولاً معادلة (4.1) على مدى جميع المجموعات التي عددها n للقيم المشاهدة، يكون لدينا:

$$\sum Y_t = n\hat{b}_0 + \hat{b}_1 \sum X_{1t} + \hat{b}_2 \sum X_{2t} + \dots + \hat{b}_k \sum X_{kt} + \sum \hat{u}_t. \quad (4.12)$$

وبفرض الشرط $\sum \hat{u}_t = 0$ ، نحصل على معادلتنا الطبيعية الأولى.

$$N1. \sum Y_t = n\hat{b}_0 + \hat{b}_1 \sum X_{1t} + \hat{b}_2 \sum X_{2t} + \dots + \hat{b}_k \sum X_{kt}.$$

نحصل على k معادلات طبيعية إضافية عن طريق ضرب المعادلة (4.11) بكل من المتغيرات المستقلة التي عددها k والجمع على مدى المجموعات التي عددها n من القيم المشاهدة. وبالتحديد،

$$\sum (X_{1t} Y_t) = \hat{b}_0 \sum X_{1t} + \hat{b}_1 \sum X_{1t}^2 + \dots + \hat{b}_k \sum (X_{1t} X_{kt}) + \sum (X_{1t} \hat{u}_t),$$

$$\sum (X_{2t} Y_t) = \hat{b}_0 \sum X_{2t} + \hat{b}_1 \sum (X_{2t} X_{1t}) + \dots + \hat{b}_k \sum (X_{2t} X_{kt}) + \sum (X_{2t} \hat{u}_t),$$

$$\sum (X_{kt} Y_t) = \hat{b}_0 \sum X_{kt} + \hat{b}_1 \sum X_{kt} X_{1t} + \dots + \hat{b}_k \sum X_{kt}^2 + \sum (X_{kt} \hat{u}_t).$$

إذا فرضنا الشروط $\Sigma(X_{it}\hat{u}_i) = 0$ لكل $(i = 1, 2, \dots, k)$ ، يسقط الحد الأخير من كل من هذه المعادلات، ونحصل، بذلك، على المعادلات الطبيعية الباقية (التي يبلغ عددها k).

$$N2. \quad \Sigma(X_{1t}Y_t) = \hat{b}_0 \Sigma X_{1t} + \hat{b}_1 \Sigma X_{1t}^2 + \dots + \hat{b}_k \Sigma(X_{1t}X_{kt})$$

$$N2. \quad \Sigma(X_{1t}Y_t) = \hat{b}_0 \Sigma X_{1t} + \hat{b}_1 \Sigma(X_{2t}X_{1t}) + \dots + \hat{b}_k \Sigma(X_{1t}X_{kt})$$

$$N(k+1). \quad \Sigma(X_{kt}Y_t) = \hat{b}_0 \Sigma X_{kt} + \hat{b}_1 \Sigma(X_{kt}X_{1t}) + \dots + \hat{b}_k \Sigma X_{kt}$$

لاحظ أن عمليات الجمع في المعادلات الطبيعية أعلاه تعتمد، فقط، على قيم المتغير التابع والمتغيرات المستقلة، فقط. وهكذا فحالما تتوافر لدينا عينة من المشاهدات فإنه يمكن حساب قيم هذه المجاميع. ويترتب على ذلك، أنه يمكن النظر إلى المعادلات الطبيعية أعلاه على أنها مجموعة مكونة من $(k+1)$ من المعادلات في $(k+1)$ من المجاهيل $\hat{b}_0, \hat{b}_1, \dots, \hat{b}_k$ وسنكون، بعامة، قادرين على تحديد قيم هذه المقدرات عن طريق حل المجموعة السابقة من المعادلات الطبيعية.

للتوضيح، اعتبر حالة الانحدار المتعدد الذي تحتوي على متغيرين مستقلين:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t.$$

سيستج، باتباع الطرق الموصوفة أعلاه، مجموعة من ثلاث معادلات طبيعية:

$$\Sigma Y_t = n\hat{b}_0 + \hat{b}_1 \Sigma X_{1t} + \hat{b}_2 \Sigma X_{2t},$$

$$\Sigma(X_{1t}Y_t) = \hat{b}_0 \Sigma X_{1t} + \hat{b}_1 \Sigma X_{1t}^2 + \hat{b}_2 \Sigma(X_{1t}X_{2t}),$$

$$\Sigma(X_{2t}Y_t) = \hat{b}_0 \Sigma X_{2t} + \hat{b}_1 \Sigma(X_{2t}X_{1t}) + \hat{b}_2 \Sigma X_{2t}^2.$$

افترض أن حساباتنا مع القيم المشاهدة Y_t ، X_{1t} و X_{2t} تعطينا:

$$n=10 \quad \Sigma X_{1t}=2 \quad \Sigma X_{2t}=2$$

$$\Sigma X_{1t}^2=6 \quad \Sigma(X_{1t}X_{2t})=1 \quad \Sigma X_{2t}^2=4$$

$$\Sigma Y_t=5 \quad \Sigma(X_{1t}Y_t)=6 \quad \Sigma(X_{2t}Y_t)=7$$

وبإدخال هذه القيم المحسوبة في المعادلات الطبيعية، نحصل على:

$$5 = 10\hat{b}_0 + 2\hat{b}_1 + 2\hat{b}_2,$$

$$6 = 2\hat{b}_0 + 6\hat{b}_1 + \hat{b}_2,$$

$$7 = 2\hat{b}_0 + \hat{b}_1 + 4\hat{b}_2.$$

ويعطينا حل هذه المجموعة من المعادلات:

$$\hat{b}_0 = 0.045, \quad \hat{b}_1 = 0.727, \quad \hat{b}_2 = 1.545.$$

ولذا، تكون معادلتنا المقدرة للانحدار هي:

$$\hat{Y} = 0.045 + 0.727X_1 + 1.545X_2.$$

ويعطينا هذا تقديرا للقيمة المتوسطة لـ Y المرتبطة مع أي مجموعة محددة من القيم لـ X_2, X_1 .

وكما قد تخمن، فإن الحل الفعلي لمجموعة المعادلات الطبيعية لتحديد القيم المقدرة للمعاملات يمكن أن يتضمن عددا ضخما من الحسابات، حتى وإن كان عدد المتغيرات صغيرا. لذا، ولأغراض عملية، عادة ما يتطلب استخدام تحليل الانحدار المتعدد استعمال الحاسوب. إلا أنه، من حيث المبدأ، من المهم أن تفهم كيف تحدد القيم المقدرة للمعاملات. وسيساعد هذا ليس، فقط، في تفسير النتائج على نحو صحيح، ولكن (وبمساعدة بعض الموضوعات التي سنعرضها فيما بعد) في اكتشاف الصعوبات.

مشكلة تعدد العلاقات الخطية التام

قلنا في الافتراض الخامس أعلاه أن منهجنا للتقدير يفشل إذا كان واحد (أو أكثر) من المتغيرات المستقلة مولفا خطيا من المتغيرات المستقلة الأخرى. ونحن الآن في وضع يسمح لنا بمعرفة السبب. افترض (على سبيل المثال) أن X_k في المعادلة (4.3) يساوي:

$$X_{kt} = c_0 + c_1 X_{1t} + c_2 X_{2t} + \dots + c_{k-1} X_{(k-1)t}, \quad (4.13)$$

حيث c_0, c_1, \dots, c_{k-1} ثوابت. وبالمناسبة، فإن بعض هذه الثوابت قد يكون صفرياً. تذكر من المناقشة السابقة أننا أوجدنا عدد $(k+1)$ من المعادلات الطبيعية، عن طريق ضرب المعادلة (4.11) بـ X_{kt} ومن ثم الجمع، ووضع $\sum (X_{kt} \hat{u}_t) = 0$ ، فإذا كان X_{kt} يساوي التعبير الموجود في المعادلة (4.13) فإنه يمكننا اشتقاق المعادلة الطبيعية رقم $(k+1)$ طريق ضرب المعادلة الطبيعية الأولى بوساطة c_0 ، والثانية في c_1 وهلم جرا، ثم جمعها بعد ذلك. فمثلاً يمكننا في ضوء المعادلة (4.13) التعبير عن الجانب الأيسر من المعادلة الطبيعية $(k+1)$ على النحو التالي:

$$\sum (Y_t X_{kt}) = c_0 \sum Y_t + c_1 \sum (Y_t X_{1t}) + \dots + c_{k-1} \sum (Y_t X_{(k-1)t}).$$

نستحث القارئ أن يقنع نفسه بهذه القاعدة من خلال العمل مع مثال لنموذج الانحدار المتعدد بثلاثة متغيرات مستقلة.

ولكن هذا يعني أن المعادلة الطبيعية $(k+1)$ ليست معادلة مستقلة، إنها توليفة خطية من المعادلات الـ (k) الأولى. في هذه الحال، نرغب في اشتقاق مقدرات لعدد $(k+1)$ معلمات: b_0, b_1, \dots, b_k ، ولكن، تتوافر لدينا، فقط، معادلات مستقلة عددها k . ونتيجة لذلك لا يمكننا (عموماً) حل هذه المعادلات للحصول على مقدرات وحيدة لهذه المعلمات. وبدقة أكثر تنحصر مشكلتنا في أنه طالما أن واحداً من متغيراتها المستقلة عادة ما يكون مجموعاً موزوناً من قيم المتغيرات المستقلة الأخرى، فلن نكون قادرين على عزل تأثيره على المتغير التابع من تأثير المتغيرات الأخرى.

ولكننا، في هذه الحال، يمكننا تقدير الآثار المجمعة لهذه المتغيرات. فمثلاً، إذا قمنا بإحلال المعادلة (4.13) في المعادلة (4.3) فإننا نحصل على:

$$\begin{aligned} Y_t &= (b_0 + b_k c_0) + (b_1 + b_k c_1) X_{1t} + \dots \\ &\quad + (b_{k-1} + b_k c_{k-1}) X_{(k-1)t} + u_t \\ &= d_0 + d_1 X_{1t} + \dots + d_{k-1} X_{(k-1)t} + u_t, \end{aligned} \quad (4.14)$$

حيث إنه (عموما) تكون $d_i = (b_i + b_k c_i)$ في المعادلة (4.14)، لدينا معلمات عددها $k: d_0, d_1, \dots, d_{k-1}$ ، فإذا لم يكن هناك علاقات أخرى من النوع الموجود في المعادلة (4.13) فإن لدينا عدد k من المتغيرات المستقلة التي تحقق افتراضنا الخامس. باختصار، يمكننا الآن اشتقاق k من المعادلات الطبيعية وحلها للوصول إلى مقدرات وحيدة لـ $\hat{d}_0, \hat{d}_1, \dots, \hat{d}_{k-1}$.

وطالما أن $d_i = (b_i + b_k c_i)$ ، فإنه لا يمكننا أن نعامل مقدرنا لـ d_i كمقدر b_i أي لن نقدر، عموما، على تقدير تأثير X_i على Y . هناك حالة استثنائية وحيدة وهي عندما تكون $c_i = 0$. فعلى سبيل المثال، إذا كانت $c_5 = 0$ فإن $d_5 = b_5$ ، حيث يمكننا معاملة مقدرنا لـ d_5 كمقدر لقيمة b_5 . وبدلالة المعادلة (4.13) يمكننا أن نجد أن قيمة معينة لـ c_i تساوي الصفر إذا كانت قيمة المتغير X_k لا تعتمد على قيمة X_i . وعموما، لا يمكننا تقدير معامل متغير مستقل منحل degenerate، مثل X_{kt} ، الذي هو توليفة خطية للمتغيرات المستقلة الأخرى في معادلة الانحدار. على سبيل المثال، إذا كان:

$$X_{2t} = 3 - 17X_{1t} + 8X_{5t},$$

حيث، فإنه، وباستثناء ($c_0 = 3$ ، $c_1 = -17$ ، $c_3 = 8$) فإن جميع c_i 's تساوي الصفر. ونتيجة لذلك، يمكن تقدير جميع الـ c_i 's الأصلية (باستثناء b_0, b_1, b_2 و b_5). باختصار، نحث القاري على ألا يحفظ هذه النتيجة عن ظهر قلب، ولكن نطلب منه تعلم الطريقة التي اشتققناها: إذا كان واحد (أو أكثر) من المتغيرات المستقلة مرتبطين خطيا ببعض المتغيرات الأخرى نستخدم، ببساطة، العلاقات الخطية للإحلال محل المتغيرات المنحلة وبعد ذلك نقدر معلمات النموذج المختزل للانحدار. ثم نحدد أخيرا معلمات نموذج الانحدار الأصلي التي سنقوم بتقديرها عن طريق مقارنة معلمات نموذج الانحدار الأصلي مع معلمات معادلة الانحدار المختزل.

(٣-٤) خصائص المقدرات واختبار الفرضيات

تفسير المقدرات*

كما أكدنا، من قبل، خلال هذا الفصل، أنه، لتقدير تأثير متغير معين مثل X_k على Y ، ينبغي علينا أن نعزل، أو نأخذ في الحسبان، تأثير المتغيرات المستقلة الأخرى كافة. وحيث، فقط، يمكننا عزل تأثير X_k على Y . وعلى الرغم من أن ذلك لم يكن واضحاً من المجموعة السابقة من معادلاتنا الطبيعية، إلا أن ذلك هو ما يفعله منهجنا في التقدير بالضبط. وقد استخدمنا في ملحق هذا الفصل، والذي نشجع القارئ بقوة على تتبعه، منهجاً بديلاً لاشتقاق صيغ مقدراتنا، \hat{b}_i . وبهذه الطريقة يمكننا معرفة الكيفية التي تستطيع من خلالها مناهجنا للتقدير فصل sorting تأثير مختلف المتغيرات المستقلة. إضافة إلى ذلك، من السهل، مع وجود الصيغ الصريحة للمقدرات في الملحق، إثبات أن هذه المتغيرات غير متحيزة، كما في حالة انحدار المتغيرين.

إلا إننا، عند هذه النقطة، نقدم مناقشة مختصرة معتمدة على القريحة لما رأيناه في الملحق. افترض أننا نحاول، باستخدام المعادلة (4.3)، تقدير تأثير تغير وحدة من X_k على Y ، أو أننا، نحاول تقدير b_k بافتراض ثبات العوامل الأخرى. نعلم أن b_k لن تكون قابلة للتقدير إذا كانت X_k توليفة خطية من المتغيرات الأخرى. افترض أن X_k ليست توليفة خطية وهذا لا يتضمن بالطبع أن X_k غير مرتبطة، مع المتغيرات المستقلة الأخرى. على سبيل المثال ففي دالة الاستهلاك قد تكون X_k الأصول السائلة و X_1 قد يكون الدخل المتاح. فإذا كان الأمر كذلك فإننا نتوقع، بالتأكيد، أن ترتبط X_k طردياً مع X_1 ، ولكن، كما هو الحال بالنسبة لطول الفرد ووزنه، هذين المتغيرين لا ينبغي أن يرتبطا ببعضهما ارتباطاً تاماً.**

* هذا البحث صعب بعض الشيء، ولذا، فإنه باستثناء فهم القاعدة (4.19)، يمكن تجاهل المعلومات الأخرى في هذا البحث عند القراءة الأولى بدون فقدان تواصل المعلومات.

** إلى حد ما، لا يحدد دخل الفرد الأصول السائلة التي يملكها تماماً، على سبيل المثال، فقد يكون هناك فردين لهما الدخل نفسه ولكن أصولهما السائلة مختلفة.

وإذا قمنا بتعيم النتيجة السابقة بعض الشيء نذكر في نموذجنا للانحدار (4.3) قد تكون X_k مرتبطة ببعض المتغيرات المستقلة أو بها كافة وإن كان هذا الارتباط ليس تاما. فإذا كان ذلك صحيحا فإنه يمكننا أن نفسر، جزئيا في الأقل، قيمة X_k بدلالة قيم المتغيرات المستقلة الأخرى. افترض أننا حاولنا عمل هذا بنموذج انحدار خطي يربط X_k بالمتغيرات المستقلة الأخرى:

$$X_{kt} = c_0 + c_1 X_{1t} + \dots + c_{k-1} X_{(k-1)t} + v_{kt}, \quad (4.15)$$

حيث u_{kt} هو الخطأ العشوائي. افترض أننا نرغب في تقدير المعلمات c_0, \dots, c_{k-1} في المعادلة (4.15) عن طريق منهج المتغيرات المساعدة* والحصول على المقدرات $\hat{c}_0, \hat{c}_1, \dots, \hat{c}_{k-1}$ باستخدام هذه المقدرات، تكون قيمتنا المفسرة (أو المحسوبة لـ X_k والمناظرة لـ X_1, \dots, X_{k-1}) هي:

$$\hat{X}_{kt} = \hat{c}_0 + \hat{c}_1 X_{1t} + \dots + \hat{c}_{k-1} X_{(k-1)t}, \quad (4.16)$$

وطالما أننا افترضنا أنه X_{kt} ليست مولفا خطيا للمتغيرات المستقلة الأخرى، فإننا نعلم أنه عموما، أن $X_{kt} = \hat{X}_{kt}$. ولذلك يمكننا أن نضع X_{kt} في الشكل:

$$X_{kt} = \hat{X}_{kt} + \hat{v}_{kt}, \quad (4.17)$$

حيث إن \hat{v}_{kt} هي ذلك الجزء من X_{kt} والذي لا يمكن أن ينسب إلى المتغيرات المستقلة الأخرى (أو $\hat{v}_{kt} = X_{kt} - \hat{X}_{kt}$) ويطلق على هذا الحد \hat{v}_{kt} الباقي "The residual" في الانحدار الذي يربط X_k بالمتغيرات المستقلة X_1, \dots, X_{k-1} .

وكما قد يخمن الفرد، فإن \hat{v}_{kt} تكون مهمة جدا في تقدير b_k ، لأنها تمثل الجزء من X_{kt} الذي يكون مستقلا، إلى حد ما، عن المتغيرات المستقلة الأخرى. فعلى سبيل المثال، من المعادلة (4.17) والمعادلة (4.16) نجد أنه، إذا كانت $\hat{v}_{kt} = 0$ ، فإننا سنعاني وجود الارتباط الخطي المتعدد التام، ولن نستطيع تقدير b_k . وفي

* نشق معادلتنا الطبيعية عن طريق وضع:

$$\sum \hat{v}_{kt} = 0, \sum (\hat{v}_{kt} X_{1t}) = 0, \dots, \sum (\hat{v}_{kt} X_{k-1t}) = 0.$$

هذه الحالة (وكما في الملحق)، أنه يمكننا أن نعبر عن المقدّر لـ b_k على النحو:

$$\hat{b}_k = \frac{\sum (\hat{v}_{kt} Y_t)}{\sum \hat{v}_{kt}^2} = \frac{\sum (X_{kt} - \hat{X}_{kt}) Y_t}{\sum (X_{kt} - \hat{X}_{kt})^2} \quad (4.18)$$

أي أن حل معادلاتنا الطبيعية لـ \hat{b}_k يمكن التعبير عنه في الشكل (4.18) لاحظ أن قيم الباقي \hat{v}_k تعتمد فقط على قيم X_1 إلى X_k ، وهكذا، يمكن مشاهدتها. وينطبق القول نفسه على مقدراتنا الأخرى. وبالتحديد، نثبت في الملحق، أن:

$$\hat{b}_i = \frac{\sum (\hat{v}_{it} Y_t)}{\sum \hat{v}_{it}^2} = \frac{\sum (X_{it} - \hat{X}_{it}) Y_t}{\sum (X_{it} - \hat{X}_{it})^2}, \quad i = 1, \dots, k, \quad (4.19)$$

حيث إن \hat{v}_{it} هو الباقي من انحدار X_i على المتغيرات المستقلة الأخرى كافة* باختصار، يعتمد مقدّرنا (\hat{b}_i) على ذلك الجزء من X_i (وهو \hat{v}_i) غير المرتبط خطياً مع المتغيرات المستقلة الأخرى. وبالمقابل يمكننا الحصول على المقدّر (\hat{b}_i) عن طريق استعمال ذلك الجزء من X_i ، فقط. الذي لا يكون مولفاً خطياً تاماً مع المتغيرات المستقلة الأخرى كافة.

والآن دع \hat{Y}_{it} هو ذلك الجزء من Y_t المرتبط خطياً تاماً مع المتغيرات المستقلة كافة باستثناء X_{it} * حينئذ، سيكون $(Y_t - \hat{Y}_{it})$ هو ذلك الجزء من Y_t غير المرتبط خطياً

* لاحظ التشابه بين صيغة \hat{b} في حالة الانحدار البسيط وحالة \hat{b}_i ، في (4.19) ففي نموذج الانحدار البسيط

$$\hat{b} = \frac{\sum (X_t - \bar{X}) Y_t}{\sum (X_t - \bar{X})^2} \quad \text{في الفصل الثاني، أوضحنا أن:}$$

وعلى سبيل تدريب القارئ، عليه أن يثبت أن صيغة \hat{b} في حالة الانحدار البسيط هي حالة خاصة من (4.19)؛ أي أنه، في حالة الانحدار البسيط، فإن $\hat{X}_t = \bar{X}$. [مساعدة في الحل: في حالة الانحدار البسيط تكون المعادلة المناظرة لـ (4.15) هي $X_t = c + v_t$.

بهذه المتغيرات المستقلة. بعد ذلك، نلاحظ أنه طالما أن \hat{Y}_{ii} في المعادلة (4.19) على النحو:

$$\sum \hat{v}_{ii} = 0, \sum (\hat{v}_i X_{ji}) = 0 \quad j \neq i, \quad (4.20)$$

فلدينا بالضرورة:

$$\sum (\hat{v}_{ii} \hat{Y}_{ii}) = 0. \quad (4.21)$$

ومن المعادلة (4.21)، وطالما أن $Y_t = \hat{Y}_{ii} + (Y_t - \hat{Y}_{ii})$ ، يمكننا التعبير عن المقدّر \hat{b}_i في المعادلة (4.19) على النحو:

$$\hat{b}_i = \frac{\sum \hat{v}_{ii} (Y_t - \hat{Y}_{ii})}{\sum \hat{v}_{ii}^2} = \frac{\sum (X_{ii} - \hat{X}_{ii})(Y_t - \hat{Y}_{ii})}{\sum (X_{ii} - \hat{X}_{ii})^2}. \quad (4.22)$$

ينبغي أن يكون واضحاً الآن كيف يعزل منهجنا تأثير المتغيرات المستقلة المختلفة على المتغير التابع Y_t . فالمقدّر \hat{b}_i في المعادلة (4.22) يعتمد، فقط، على قيم X_{ii} و Y_t بعد استبعاد التأثير الخطي لجميع المتغيرات المستقلة الأخرى على كلا المتغيرين. في الحقيقة، يمكننا أن نتخيل أن منهجنا الحالي في التقدير يحول حالة المتغيرات المتعددة إلى حالة المتغيرين عن طريق استبعاد (طرح) تأثير المتغيرات الأخرى.

تباينات المقدرات

ينبغي عليك النظر لكيفية اشتقاق المعادلة (4.19) في ملحق هذا الفصل خاصة وأن المعالجة الدقيقة هناك تمكننا أولاً: من إثبات أن \hat{b}_i مقدر غير

* يمكن أن نحصل على ذلك عن طريق انحدار Y_t على المتغيرات المستقلة كافة باستثناء X_{ii} ، ثم حساب قيمته المتوقعة. على سبيل المثال، سنعتبر نموذج الانحدار:

$$Y_t = \gamma_0 + \gamma_1 X_{1t} + \dots + \gamma_{i-1} X_{(i-1)t} + \gamma_{i+1} X_{(i+1)t} + \dots + \gamma_k X_{kt} + w_t,$$

حيث w_t هو الخطأ العشوائي. بعدئذ، نستخدم طريقة المتغير المساعد للحصول على:

$$\hat{\gamma}_0, \hat{\gamma}_1, \dots, \hat{\gamma}_{(i-1)}, \hat{\gamma}_{(i+1)}, \dots, \hat{\gamma}_k$$

وحيث نحصل على:

$$\hat{Y}_{ii} = \hat{\gamma}_0 + \hat{\gamma}_1 X_{1i} + \dots + \hat{\gamma}_{i-1} X_{(i-1)i} + \hat{\gamma}_{i+1} X_{(i+1)i} + \dots + \hat{\gamma}_k X_{ki}.$$

متحيز لـ b_i (أي أن $E(\hat{b}_i) = b_i$). ثانيا: إيجاد صيغ للتباينات الشرطية لـ \hat{b}_i .
ويمكننا، بالتحديد، أن نثبت:

$$\text{var}(\hat{b}_i) = \sigma_u^2 = \frac{\sigma_u^2}{\sum \hat{v}_{ii}^2}, \quad \text{for } i = 1, 2, \dots, k, \quad (4.23)$$

حيث σ_u^2 هو تباين الخطأ العشوائي في معادلة الانحدار المتعدد الأصلية (4.3). *

فترات الثقة واختبار الفرضيات: بعض المقدمات

بالاستمرار في افتراض، أن الأخطاء العشوائية موزعة توزيعاً طبيعياً كما في حالة الانحدار البسيط، فسيكون صحيحاً أيضاً، - بافتراض قيم معطاة للمتغيرات المستقلة - أن المقدرات \hat{b}_i كافة ستكون بدورها موزعة توزيعاً طبيعياً. ويتبع ذلك من كون \hat{b}_i توليفة خطية من الأخطاء العشوائية. وكما ناقشنا في الفصل الثالث فإن التوليفات الخطية من المتغيرات الموزعة توزيعاً طبيعياً تكون نفسها موزعة توزيعاً طبيعياً. وهذا يمكننا من تجميع نتائجنا رمزياً على النحو: **

$$\hat{b}_i \text{ is } N(b_i, \sigma_{\hat{b}_i}^2), \quad i = 0, 1, \dots, k, \quad (4.24)$$

حيث:

* لاحظ، مرة أخرى، التشابه بين صيغ التباين لـ \hat{b} في حالة الانحدار ذي المتغيرين و \hat{b}_i في حالة الانحدار المتعدد. والفرق، مرة أخرى، هو أن الحد $\Sigma(X_i - \bar{X})^2$ في مقام صيغة التباين لـ \hat{b} في حالة انحدار المتغيرين قد تم إحلاله بـ $\Sigma(X_{ii} - \bar{X}_{ii})^2$.

** المتغير المستقل المناظر لـ b_0 هو X_{0i} ، $i = 1, 2, \dots, n$. لذلك، سيكون \hat{v}_{0i} هو الباقي من انحدار X_{0i} على المتغيرات المستقلة الأخرى كافة X_{1i}, \dots, X_{ki} .

$$X_{0i} = e_1 X_{1i} + e_2 X_{2i} + \dots + e_k X_{ki} + v_{0i}$$

لاحظ أن معادلة الانحدار هذه لا تشمل على حد ثابت، لأن الحد الثابت هو المتغير التابع. وعلى العكس، فإن معادلات الانحدار التي تعرف الأخطاء العشوائية v_{ii} مثل المعادلة (4.15) تحتوي على الحدود الثابتة، وذلك بسبب أن واحداً من المتغيرات المفسرة هو المتغير التابع. باختصار، فإن X_{0i} ينبغي أن ينظر إليه على أنه متغير مفسر آخر.

$$\sigma_{\hat{b}_i}^2 = \frac{\sigma_u^2}{\sum \hat{v}_{ii}^2}$$

وتبقى الصعوبة أمامنا في أنه لا يمكن تحديد تباينات مقدراتنا لأن σ_u^2 غير معلومة. ولذا، نحتاج لبناء مقدر لـ σ_u^2 من عيئتنا من المشاهدات. وطالما أن σ_u^2 هي القيمة المتوسطة لـ u_i^2 :

$$\sigma_u^2 = E[u_i^2].$$

يكون من المنطقي، حيثئذ، إذا علمنا قيم b_i ، أن نأخذ متوسط تباينات العينة كمقدر لـ σ_u^2 :

$$\frac{\sum u_i^2}{n} = \frac{\sum (Y_i - b_0 - b_1 X_{1i} - \dots - b_k X_{ki})^2}{n}. \quad (4.25)$$

ولسوء الحظ، فإننا لانعلم قيم المعلامات، إلا أنه لدينا مقدرات لها. ومن ثم، يمكننا إحلال كل من b_i في المعادلة (4.25) بمقدره، ومن ثم، الحصول على مقدر لتباين الخطأ العشوائي.

$$\hat{\sigma}_u^2 = \frac{\sum_{i=1}^n (Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \dots - \hat{b}_k X_{ki})^2}{n - (k+1)} \quad (4.26)$$

لاحظ أن المقام في المعادلة (4.26) هو $[n - (k+1)]$ ويعكس هذا خسارة عدد $(k+1)$ من درجات الحرية في البسط نتيجة تقدير معلامات عددها $(k+1)$. وكما في حال انحدار المتغيرين، يكون صحيحاً، أيضاً:

$$E(\hat{\sigma}_u^2) = \sigma_u^2.$$

أي أن مقدرنا في المعادلة (4.26) غير متحيز. ويمكننا باستخدام المعادلة (4.26)، تقدير تباين كل من \hat{b}_i بوساطة:

$$\hat{\sigma}_{\hat{b}_i}^2 = \frac{\hat{\sigma}_u^2}{\sum \hat{v}_{ii}^2}. \quad (4.27)$$

فترات الثقة واختبار الفرضيات

نحن الآن في وضع يمكننا من بناء فترات الثقة للمعاملات المفردة. وباستخدام نتائجنا من الفصل الثالث، نلاحظ أنه طالما أن \hat{b}_i متغير طبيعي $[\hat{b}_i \sim N(b_i, \sigma_{\hat{b}_i}^2)]$ فسيكون:

$$\frac{\hat{b}_i - b_i}{\sigma_{\hat{b}_i}}$$

متغيرا طبيعيا معياريا $N(0,1)$ ، حيثئذ وبالطريقة نفسها التي اتبعناها في حالة الانحدار البسيط يمكننا إنشاء فترات الثقة، واختبار الفرضيات بدلالة المنحنى الطبيعي إذا كان σ_u^2 (ومن ثم $\sigma_{\hat{b}_i}^2$) معلوما. على سبيل المثال ينبغي أن يثبت القارئ لنفسه أن فترة الثقة 95% لـ b_i ، مبنية على المنحنى الطبيعي، هي:

$$\hat{b}_i \pm 1.96 \sigma_{\hat{b}_i}.$$

وكما في حالة انحدار المتغيرين، تكون u^2 (عموما) غير معلومة، ولذا ينبغي تقديرها. ونتيجة لذلك ننشئ فترات الثقة أو نوجد نسب t لاختبار الفرضيات بدلالة توزيع t . وللقيام بذلك نلاحظ أن:

$$\frac{\hat{b}_i - b_i}{\hat{\sigma}_{\hat{b}_i}} \quad (4.28)$$

هو متغير t بدرجات حرية $(n-k-1)$. لاحظ أن درجات الحرية للمتغير t في المعادلة (4.28) تكون مساوية دائما مقام مقدر التباين في المعادلة (4.26). وعلى سبيل المقارنة تذكر أنه، في حالة الانحدار البسيط، تكون $k=1$ ولذا (مع وجود معلمتين ينبغي تقديرهما يكون لدينا توزيع t بدرجات حرية قدرها $(n-2)$).

يمكننا الآن، باستخدام المعادلة (4.28)، أن نتبع المنهج نفسه الذي طورناه في الفصل الثالث لاختبار الفرضيات المرتبطة بأي من المعاملات المفردة. على سبيل المثال، افترض أنه يوجد لدينا النموذج التالي:

$$Y_t = b_0 + b_1 X_{1t} + \dots + b_9 X_{9t} + u_t, \quad t=1, \dots, 25, \quad (4.29)$$

ونرغب في اختبار الفرضيات (عند 0.05 حجم النوع الأول من الخطأ):

$$H_0: b_3 = 0,$$

$$H_1: b_3 \neq 0,$$

وتتوافر لدينا عينة مكونة من 25 مشاهدة.

في هذه المسألة يكون لدينا $n=25$ و $k=9$ ومن ثم:

$$\frac{\hat{b}_3 - b_3}{\hat{\sigma}_{\hat{b}_3}}$$

يكون متغير t بدرجات حرية عددها $(25-9=15)$. ومن الجدول الإحصائي ٢ لتوزيع t نجد أن فترة الثقة 95% لـ b_3 هي:

$$(\hat{b}_3 \pm 2.131 \hat{\sigma}_{\hat{b}_3}). \quad (4.30)$$

وبالاستمرار، كما فعلنا في الفصل الثالث، يمكننا استخدام عينتنا لتقويم $\hat{b}_3, \hat{\sigma}_{\hat{b}_3}$ وبالتعويض عن هذه القيم في المعادلة (4.36)، ومعرفة ما إذا كانت الفترة الناتجة تغطي القيمة المفترضة الصفرية أم لا. فإذا كانت تغطيها فسوف نقبل H_0 ، أما إذا لم تكن تغطيها فسوف نرفضه. وبالمقابل يمكننا، ببساطة، أن نقيم نسبة t واتباع قاعدتنا التجريبية، نقارن نسبة t المطلقة مع ٢. وفي حالتنا هذه تكون القيمة الحرجة الدقيقة هي 2.131، وعلى أي حال ينبغي أن يكون واضحاً أنه عند اللحظة التي نعطي فيها نتيجة مثل المعادلة (4.28) تُحلُّ المشاكل المرتبطة باختبار فترات الثقة وتكوينها في حالة الانحدار المتعدد بالطريقة نفسها التي استخدمناها في حال انحدار المتغيرين.

(٤-٤) معامل التحديد المتعدد

بينا في المباحث السابقة طرق تقدير الفرضيات المرتبطة بالمعاملات الفردية في نموذج الانحدار المتعدد واختبارها. تبقى قضية إضافية متعلقة بالقوة التفسيرية لمعادلة الانحدار ككل. ما المقدار الذي يمكن تفسيره من التغير في المتغير التابع عند أخذ المتغيرات المستقلة كافة معاً؟

R^2 لحالة الانحدار المتعدد

أوجدنا في الفصل الثاني مقياس R^2 لحالة الانحدار البسيط. تذكر أن R^2 تحتوي على قيمة بين الصفر والواحد الصحيح تشير إلى جزء التغير في Y الذي يمكن أن تفسره معادلة الانحدار المقدرة. واتباع المنهج نفسه فسوف نثبت أن القاعدة نفسها R^2 بالتفسير نفسه قابلة للتطبيق، أيضا، لحالة الانحدار المتعدد.

اعتبر نموذجنا الأساسي للانحدار المتعدد التالي:

$$Y_i = b_0 + b_1 X_{1i} + \dots + b_k X_{ki} + u_i. \quad (4.31)$$

لاحظنا سابقا أنه إذا قدرت المعادلة (4.31) بوساطة طريقة المتغير المساعد فإنه يمكن التعبير عن Y_i على النحو:

$$Y_i = \hat{Y}_i + \hat{u}_i, \quad (4.32)$$

حيث:

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_{1i} + \dots + \hat{b}_k X_{ki}, \quad (4.33)$$

و \hat{u}_i حيث تكون:

$$\sum \hat{u}_i = 0, \quad \sum (\hat{u}_i X_{ii}) = 0, \quad i = 1, \dots, k \quad (4.34)$$

والآن نجمع المعادلة (4.32) على مدى عيئنا لنحصل على:

$$\sum Y_i = \sum \hat{Y}_i + \sum \hat{u}_i. \quad (4.35)$$

ولما كانت $\sum \hat{u}_i = 0$ ، يكون لدينا كما في حالة الانحدار البسيط

$$\sum Y_i = \sum \hat{Y}_i. \quad (4.36)$$

وبقسمة حدود المعادلة (4.36) على n يكون متوسط العينة لـ Y_i يساوي متوسط العينة لـ \hat{Y}_i :

$$\bar{Y} = \bar{\hat{Y}}. \quad (4.37)$$

وبالعودة إلى المعادلة (4.32) وتربيع جانبي المعادلة:

$$Y_i^2 = \hat{Y}_i^2 + 2\hat{Y}_i \hat{u}_i. \quad (4.38)$$

وبالجمع على مدى العينة، نحصل على:

$$\sum Y_i^2 = \sum \hat{Y}_i^2 + \sum \hat{u}_i^2 + 2 \sum (\hat{Y}_i \hat{u}_i). \quad (4.39)$$

ويساوي الحد الأخير في المعادلة (4.39) الصفر، ولنرى ذلك لاحظ أن:

$$\begin{aligned} \sum (\hat{u}_i \hat{Y}_i) &= \sum \hat{u}_i (\hat{b}_0 + \hat{b}_1 X_{i1} + \dots + \hat{b}_k X_{ik}) \\ &= \hat{b}_0 \sum \hat{u}_i + \hat{b}_1 \sum (\hat{u}_i X_{i1}) + \dots + \hat{b}_k \sum (\hat{u}_i X_{ik}) = 0, \end{aligned}$$

ويمكن في ضوء المعادلة (4.34) تبسيط المعادلة (4.39) إلى:

$$\sum Y_i^2 = \sum \hat{Y}_i^2 + \sum \hat{u}_i^2. \quad (4.40)$$

نطرح بعد ذلك $n\bar{Y}^2$ من جانبي المعادلة (4.40):

$$\sum Y_i^2 - n\bar{Y}^2 = (\sum \hat{Y}_i^2 - n\bar{Y}^2) + \sum \hat{u}_i^2. \quad (4.41)$$

وبتذكر أنه من (4.31) أن $\bar{Y} = \widehat{\bar{Y}}$ ، يمكننا أن نكتب المعادلة (4.41) على النحو: *

$$\sum (Y_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y})^2 + \sum \hat{u}_i^2, \quad (4.42)$$

والتي تكون متماثلة مع المعادلة المناظرة لحالة الانحدار البسيط والتي تم اشتقاقها في الفصل الثاني.

تذكر أننا وضعنا هذه العلاقة على النحو:

$$TSS = RSS + ESS \quad (4.43)$$

حيث إن $TSS = \sum (Y_i - \bar{Y})^2$ و $RSS = \sum (\hat{Y}_i - \bar{Y})^2$ وأخيراً $ESS = \sum \hat{u}_i^2$. ويعبر

المجموع الكلي للمربعات TSS عن التغير في المتغير التابع حول الوسط الحسابي للعينة الذي نحاول تفسيره بمعادلتنا للانحدار. أي أنه يفترض أن يخبرنا نموذجنا للانحدار لماذا تكون Y_i غير ثابتة. وذلك الجزء الذي لا يمكن للمعادلة أن تفسره هو مجموع مربعات الخطأ (ESS). والفرق بين TSS و ESS ينبغي أن يكون ذلك الجزء الذي تفسره معادلة الانحدار وهو (RSS) مجموع مربعات الانحدار.

* نستخدم هنا إحدى قواعد الفصل الأول التي تعني أنه بالنسبة لأي متغير Z_i :

$$\sum (Z_i - \bar{Z})^2 = \sum Z_i^2 - n\bar{Z}^2.$$

وكما هو في حالة نموذج الانحدار ذي المتغيرين، نستخدم الجزء من المتغير في القيم المشاهدة لـ Y الذي يمكن أن ينسب إلى معادلة الانحدار المقدرة بوصفه مقياساً للقوة التفسيرية لمعادلة الانحدار.

$$R^2 = 1 - \frac{ESS}{TSS} = \frac{RSS}{TSS} = \frac{\sum (\hat{Y}_t - \bar{Y})^2}{\sum (Y_t - \bar{Y})^2}, \quad (4.44)$$

حيث يطلق على R^2 معامل التحديد المتعدد.

باختصار، إذا كان لدينا توفيق تام فإن القيمة المحسوبة تتساوى في كل حالة مع القيم المشاهدة ($Y_t = \hat{Y}_t$) أي أن \hat{u}_t تساوي الصفر وأن $ESS=0$ ، لذلك يكون $RSS=TSS$ ، ويحقق R^2 قيمته العظمى بأن تساوي الواحد الصحيح. وبالمقابل، إذا كانت المعادلة المقدرة لا تفسر أيًا من التغير في المتغير التابع فسيكون ESS أكبر مما يمكن ويتعادل هنا مع TSS ($ESS=TSS$). ولذا، يكون $RSS=0$. وهنا يكون $R^2 = 0$ وكلما اقترب R^2 من الواحد ازدادت القوة التفسيرية لمعادلة الانحدار المقدرة. وأخيراً، وعلى الرغم من أننا لن نعطي إثباتاً (طالما أن الحالة متماثلة مع حالة انحدار المتغيرين) نشير إلى أن R قد لا تعني أكثر من مجرد مقدر لمعامل الارتباط بين Y_t و \hat{Y}_t .

تعليق على R^2

لما كان $R^2 = 1 - ESS/TSS$ معامل التحديد، فإنه يعتمد بوضوح على كل من ESS و TSS . ومن بين هذين المحددين لـ R^2 ، ترتبط ESS بالقوة التفسيرية للنموذج. ذلك أن TSS ، ببساطة، هو مقياس لتغير العينة في المتغير الذي نحاول تفسيره، وهو بذلك لا يعتمد على سمات النموذج المستخدم لتفسير Y_t . وهكذا يرتبط R^2 بالقوة التفسيرية للنموذج بحكم ارتباط ESS بها.

اعتبر الآن الحالة التي يرغب فيها اثنان من الباحثين في تفسير قيم عددها n من Y_t ، $t=1,2,\dots,n$. افترض أن الباحث الأول يستخدم النموذج (4.31)، أما الباحث الثاني فيستخدم النموذج التالي:

$$Y_t = b_0 + b_1 X_{1t} + \dots + b_k X_{kt} + b_{k+1} X_{k+1,t} + \dots + b_{k+r} X_{k+r,t} + u_t. \quad (4.45)$$

أي أن الباحث الثاني يدخل في نموذج المتغيرات المستقلة كافة التي اعتبرها الباحث الأول، إضافة إلى بعض المتغيرات الأخرى (أي $X_{k+1,t}, \dots, X_{k+r,t}$). دع ESS_1 و RSS_2 هما مجموع مربعات الخطأ التي حصل عليها بواسطة الباحث الأول والباحث الثاني على الترتيب. حينئذ، وكما سنبين أدناه، فإن:

$$ESS_2 \leq ESS_1, \quad (4.46)$$

ونتساءل الآن هل المتغيرات الإضافية التي اعتبرها الباحث الثاني ملائمة! دع R_1^2 و R_2^2 وهما معاملي التحديد اللذين حصل عليهما الباحثان الأول والثاني على الترتيب. وبما أن $R_1^2 = 1 - ESS_1 / TSS$ و $R_2^2 = 1 - ESS_2 / TSS$ فإنه ينتج من المعادلة (4.46) أن تكون:

$$R_1^2 \leq R_2^2. \quad (4.47)$$

مرة أخرى، تظل المعادلة (4.47) صحيحة سواء كانت المتغيرات المستقلة الإضافية ملائمة أم لا. ويظل عدم التساوي في المعادلة (4.46) والمعادلة (4.47) صحيحا دائما. ولكننا نلاحظ في التطبيقات العملية، عادة، أنه إذا لم يكن $ESS_1 = 0$ و $R_1^2 = 1$ فإن:

$$ESS_2 \leq ESS_1, \text{ و } R_2^2 > R_1^2 \quad (4.48)$$

وتشير هذه النتيجة إلى أن إحصائية R^2 تستمر في التزايد ولن تنخفض أبدا بإضافة متغيرات مستقلة إضافية إلى النموذج سواء كانت هذه المتغيرات ملائمة أم لا. والسبب في ذلك هو، وكما أوضحنا في ملحق الفصل الثاني، أن مقدراتنا في حالة الانحدار ذي المتغيرين والناجمة عن استخدام طريقة المتغيرات المساعدة متطابقة مع المقدرات الناتجة باستخدام طريقة المربعات الصغرى. ويكون هذا التطابق صحيحا، أيضا، في حالة الانحدار المتعدد. أي أن مقدراتنا للمعاملات الناتجة عن استخدام طريقة المتغيرات المساعدة لنموذج الانحدار المتعدد الخطي متطابقة مع المقدرات التي يحصل عليها باستخدام طريقة المربعات الصغرى.

دع $\hat{b}_0, \dots, \hat{b}_k$ هي تقديرات معاملات الانحدار التي حصل عليها الباحث الأول. حيثئذ تكون هذه التقديرات بشكل يكون معه مجموع مربعات الخطأ:

$$ESS_1 = \sum_{t=1}^N (Y_t - \hat{b}_0 - \hat{b}_1 X_{1t} - \dots - \hat{b}_k X_{kt})^2 \quad (4.49)$$

صغيرا بقدر الإمكان. ويتبع هذا من التطابق بين طريقتي المتغير المساعد والمربعات الصغرى.

دع الآن $\bar{b}_0, \dots, \bar{b}_k, \bar{b}_{k+1}, \dots, \bar{b}_{k+r}$ هي التقديرات التي حصل عليها الباحث الثاني. حيثئذ ومرة أخرى وبسبب التماثل مع طريقة المربعات الصغرى، تكون هذه التقديرات بشكل يصبح معه ESS_2 صغيرا بقدر الإمكان عندما يكون:

$$ESS_2 = \sum_{t=1}^N (Y_t - \bar{b}_0 - \dots - \bar{b}_k X_{kt} - \bar{b}_{k+1} X_{k+1,t} - \dots - \bar{b}_{k+r} X_{k+r,t})^2. \quad (4.50)$$

وفي هذه الحال، سيتساوى ESS_2 مع ESS_1 . هو أن إحدى المجموعات الممكنة من القيم لـ هي:

$$\bar{b}_i = \hat{b}_i, \quad i=0, \dots, k \quad (4.51)$$

$$\bar{b}_{k+j} = 0, \quad j=1, \dots, r.$$

وبسبب أن ESS_2 لن يزيدا ابدا على ESS_1 ، ينتج عن ذلك أن القوة التفسيرية للنموذج الثاني (كما تقاس بوساطة R^2) لن تكون اقل من نظيرتها في النموذج الأول ابدا. ولكن يحصل على قيم أخرى لـ $\bar{b}_0, \dots, \bar{b}_k$ ولذا يكون $ESS_2 < ESS_1$ ومن ثم، يكون $R_2^2 > R_1^2$.

معامل التحديد المعدل \bar{R}^2

يرغب الباحثون، عادة، في مقارنة نماذجهم بدلالة مختلف المقاييس لمدى جودة "goodness" النموذج. وتوحي مناقشتنا أن إحصائية R^2 يمكن أن تكون واحدة من هذه المقاييس ولكن مع الحذر. ذلك أنه إذا كانت R^2 تزداد، عادة، بإضافة

متغيرات مستقلة جديدة في النموذج (بغض النظر عن مدى ملائمتها له)، فإن الباحث يمكنه الوصول إلى نموذج ممتاز يتفوق على النماذج الأخرى باستخدام هذا المقياس عن طريق إضافة متغيرات مستقلة جديدة.

وتكمن الصعوبة التي نواجهها مع إحصائية R^2 في عدم وجود جزاء مترتب على زيادة عدد المتغيرات المستقلة المستخدمة في تفسير المتغير التابع. ولذا، نستخدم في بعض الأحيان شكلاً معدلاً من R^2 . ويتضمن هذا الشكل الجزاء المشار إليه أعلاه.

يرمز، عادة، لمعامل التحديد المعدل بالرمز \bar{R}^2 . عموماً افترض أننا نأخذ نموذج الانحدار الخطي بحد ثابت، وبعدد p من معاملات الانحدار كعدد إجمالي، فعلى سبيل المثال، بالنسبة للمعادلة (4.31) تكون $p=k+1$ وبالنسبة للمعادلة (4.45) تكون $p=k+r+1$ حيث، يكون معامل التحديد المعدل هو:

$$\bar{R}^2 = 1 - \frac{ESS/(n-p)}{TSS/(n-1)} \quad (4.52)$$

حيث إن n هي حجم العينة، ESS و TSS هما مجموع مربعات الخطأ ومجموع المربعات الكلية للنموذج موضع الاعتبار على الترتيب.*

ويمكن أن تزداد قيمة \bar{R}^2 ، أو أن تظل على حالها، أو تنقص إذا أضفنا متغيرات مستقلة إضافية للنموذج، وسبب ذلك هو أن مثل هذه المتغيرات المستقلة الإضافية سوف تخفض عادة من قيمة ESS ، ولكنها سوف تقلل، أيضاً، من $n-p$. ولذا لا يمكن التنبؤ باتجاه التغير في $ESS/(n-p)$ ، ومن ثم، اتجاه التغير في \bar{R}^2 .

ويشعر الباحثون، غالباً، أنه إذا أضيف متغير مهم أو مجموعة متغيرات مهمة للنموذج فإن النقص في ESS سيفوق الحجم اللازم للتعويض عن النقص في $(n-p)$ ولذا سترداد \bar{R}^2 . بالطبع، ستكون هذه النتيجة صحيحة لبعض المتغيرات

* باستخدام المعادلة (4.52) والمعادلة (4.44)، يمكن إثبات أن $\bar{R}^2 \leq R^2$ إذا كانت $R^2 = 1$ ، حيث، فإن $\bar{R}^2 = R^2$ ، ولكن، إذا كانت $R^2 < 2$ و $p \geq 2$ فإن $\bar{R}^2 < R^2$ للمهتمين بهذا الموضوع، نقدم إثباتاً لهذه النتائج في الملحق B في ختام هذا الفصل

المهمة ولكننا نؤكد هنا أن هذه ليست نتيجة نظرية دقيقة لأنه يمكن إيجاد أمثلة مناقضة لها. فعلى سبيل المثال في المعادلة (4.45)، افترض أن $r=1$ وأن $X_{k+1,t}$ متغير مهم لأن معامل لايساوي الصفر أي $b_{k+1} \neq 0$. في هذه الحال، يعتمد المتغير التابع Y_t ، في الحقيقة، على $X_{k+1,t}$ (بالإضافة إلى المتغيرات الأخرى). على الرغم من هذه الأهمية، يمكن أن يظل $\bar{R}_1^2 < \bar{R}_2^2$ حيث إن \bar{R}_1^2 هو معامل التحديد المعدل المناظر للمعادلة (4.45) وأن \bar{R}_1^2 المناظر للمعادلة (4.31).

ولما كان كثير من الباحثين يشعر أن \bar{R}^2 سيتزايد إذا أضيفت متغيرات «مهمة» وينقص إذا أضيفت متغيرات «غير مهمة» فإن إحصائية \bar{R}^2 تستخدم للمقارنة بين النماذج افترض، على سبيل المثال، أن النماذج موضع الاعتبار هي المعادلة (4.45) والمعادلة (4.31). حيثئذ، سيأخذ بعض الباحثين النموذج «الحقيقي» أو «الأفضل» على أنه ذلك النموذج الذي يحتوي على \bar{R}^2 أكبر.

وعلى الرغم من الاغراءات الحدسية لهذا النموذج، إلا أنه لا يستند على أسس علمية، ولذا، فإننا لانوصي باستخدامه. وسبب ذلك هو أن مثل هذه المقارنات هي شكل من اشكال اختيار صدق أحد النماذج أو صحته إزاء النموذج الآخر، إلا أن خصائص هذا الاختبار، كأخطائه من النوع الأول ومن النوع الثاني، غير معلومة. لذلك، وفي الواقع، فإن الباحثين الذين يقيمون النماذج بدلالة إحصائية \bar{R}^2 بهذه الطريقة المباشرة إنما يستخدمون اختبارا لم تعرف خصائصه بعد.

أما منهج الاختبار الملائم لتقويم المعادلة (4.45) بالنسبة للمعادلة (4.31) فيكون على النحو التالي: أولا، اعتبر $r=1$. في هذه الحال، يكون اختبار فرضية العدم الملائم $H_0: b_{k+1} = 0$. ومن الواضح أنه إذا كانت $b_{k+1} = 0$ فإن المعادلة (4.31) تكون محددة تحديدا صحيحا، وإذا كانت $b_{k+1} \neq 0$ غير محددة تحديدا صحيحا بينما تكون المعادلة (4.45) محددة بصورة صحيحة. افترض لتبسيط الطرح، أن الفرضية البديلة هي $H_1: b_{k+1} \neq 0$. افترض، أخيرا، أننا نرغب في جعل الخطأ من النوع الأول مساويا لـ 0.05 حيثئذ، مع تحقق افتراضات النموذج الموجودة في المبحث (٤-١)،

سوف يتم رفض H_0 لصالح H_1 إذا كانت $\left| \hat{b}_{k+1} / \hat{\sigma}_{\hat{b}_{k+1}} \right| > t_{n-k-2} 0.975$ حيث إن \hat{b}_{k+1} هو مقدرنا لـ b_{k+1} وفقا لطريقة المتغير المساعد، و $\hat{\sigma}_{\hat{b}_{k+1}}$ هو المقدر المناظر للانحراف المعياري لـ \hat{b}_{k+1} ، وقد ناقشنا طريقة هذا الاختبار في المبحث (٤-٣).
اعتبر الآن $r > 1$ ، في هذه الحال، تكون فرضية العدم موضع الاهتمام هي:

$$H_0: b_{k+1} = b_{k+2} = \dots = b_{k+r} = 0. \quad (4.53)$$

من الواضح أنه إذا كانت H_0 في المعادلة (4.53) صحيحة، فإن النموذج (4.31) يكون محددا تحديدا صحيحا. ولقد شرحت إحدى الطرق لاختبار الفرضيات من النوع الموجود في المعادلة (4.53) مقابل الفرضية البديلة ذات الطرفين في الملحق B من الفصل الخامس. مرة أخرى، وعلى العكس من اختبار \bar{R}^2 يمكن من خلال الطريقة -الموجودة في الملحق B من الفصل الخامس، وضع حجم الخطأ من النوع الأول عند المستوى المرغوب فيه (مثلا 0.05). خلاصة القول هو أنه لا ينبغي تفضيل نموذج على آخر بسبب أنه يتضمن قيمة أعلى لـ \bar{R}^2 .

وقد يختار القارئ في السبب الذي فعلناه لتضمين المناقشة الخاصة بمعامل التحديد المعدل أو \bar{R}^2 في الوقت الذي نوصي بقوة بعدم استخدامه. ان المنطق وراء ذلك مزدوج. أولا، كثيرا ما تطبع برامج الحاسوب قيم R^2 و \bar{R}^2 . فإذا أهملنا \bar{R}^2 فقد تتعجب ماذا تعني هذه الإحصائية. وثانيا، وكما أشرنا، فإن بعض الباحثين يقارنون مدى جودة التوفيق لنماذجهم من خلال مقارنة قيم \bar{R}^2 . ونحث مرة، أخرى، على الحذر من قبول مثل هذا الاختبار، لأنه، كما أشرنا من قبل، لا يستند إلى أسس علمية صحيحة.

وهناك نقطة أخيرة هي أننا نلاحظ أن هناك منهجا علميا لاختبار الفرضيات من النوع الموجود في المعادلة (4.53) المبني على مقارنات للإحصائية \bar{R}^2 ، وفي الواقع، فإن هناك اختبار آخر مبني على مقارنة إحصائيات R^2 . ولكن هذين المنهجين، في الحقيقة، يمثلان المنهج الشائع الاستخدام والموصوف في الملحق B من الفصل الخامس. ولما كانت الاعباء الحسابية لهذه الاختبارات ليست أقل من الاختبارات

الموصوفة في الملحق B، وطالما أنها لاتقدم لنا إضافات جديدة فإننا لن نشرحها في هذا الكتاب.

(٤-٥) تحليل الانحدار المتعدد: مثالان توضيحيان

بعد أن درسنا مبادئ تحليل الانحدار، سنفحص الآن دراستين فعليتين للانحدار المتعدد لتتعرف على المنهج المستخدم ولنفسر النتائج الخاصة بهاتين الدراستين.

دالة استهلاك متعددة المتغيرات

افترضنا، في بداية هذا الفصل، أن حجم الإنفاق الاستهلاكي قد يعتمد على عدد من المتغيرات إضافة إلى مستوى الدخل الحالي. وفي الحقيقة، كون الاقتصاديون عددا من نظريات الاستهلاك وأجروا اختبارات قياسية مكثفة لهذه النظريات. فعلى سبيل المثال، طرح البرت اندو - وفرانكو مودجلياني نظرية دورة الحياة للاستهلاك حيث يريان أن مستوى الاستهلاك الحالي لأحد الأفراد يعتمد على القيمة المتوقعة لتيار دخله المستقبلي مدى الحياة.* وعلى سبيل الاختبار العملي البسيط لهذا الافتراض، اقترحا أن:

$$C_t = b_0 + b_1 Y_{dt} + b_2 A_{t-1} + u_t, \quad (4.54)$$

حيث يستعمل دخل العمل المتاح الحالي (Y_{dt}) متغيرا تقريبا عن الدخل المتوقع من خدمات العمل، و A_{t-1} هو القيمة الصافية لثروة المستهلكين في نهاية الفترة ($t-1$) مقياس لدخل الملكية المتوقع. أي أن المستهلكين يبدأون الفترة t بثروة صافية قدرها A_{t-1} والتي تزودهم بالدخل الربعي أو الفائدة. قدر أندو ومودجلياني بعد ذلك قيم المعاملات في المعادلة (4.54) باستخدام تحليل الانحدار المتعدد وباستخدام بيانات سنوية عن الإنفاق الاستهلاكي، دخل العمل المتاح، وصافي الثروة، وقيست جميعا بـ ١٩٢٩-١٩٥٩ م (بعد استبعاد سنوات

* انظر: "The Life Cycle Hypothesis of Saving : Aggregate Implications and Tests." *American Economic Review*, 53(March 1963), pp. 55-84.

الحرب ١٩٤١-١٩٤٥م). وجمعا ٢٥ مشاهدة مشتركة لهذه المتغيرات، ثم استخدمنا بيانات هذه السلسلة الزمنية في نموذج الانحدار المتعدد لتقدير المعادلة (4.54) ووجدنا أن:

$$\hat{C}_t = 5.33 + 0.767Y_{dt} + 0.047A_{t-1} \quad N = 25 \quad (4.55)$$

$$(1.46) \quad (0.047) \quad (0.010) \quad R^2 = 0.999$$

(ملاحظة: الأرقام الموجودة داخل الأقواس أسفل المعاملات المقدرة هي قيم الأخطاء المعيارية المناظرة).

من أول الأشياء التي نلاحظها في نتائج اندو - مودجلياني هي أن المعاملات المقدرة، وكما توقعنا موجبة، وأن الميل الحدي للاستهلاك من دخل العمل المتاح يقع بين الصفر والواحد الصحيح، لذلك تبدو المعادلة المقدرة متسقة مع السمات المفترضة للسلوك الاستهلاكي.

دعنا نلاحظ، بعد ذلك، من المعادلة (4.55) أنه - في الحالات الثلاث جميعا - تزيد تقديرات المعلمة على ثلاثة أضعاف تقديرات الأخطاء العشوائية. وباستخدام قاعدتنا التجريبية المرتبطة بنسب t ، يتضمن هذا أنه إذا اعتبرت أية فرضية من فرضيات العدم $b_2 = 0$ و $b_1 = 0$ و $b_0 = 0$ مقابل الفرضيات البديلة ذات الذيل الواحد أو ذات الذيلين فسوف ترفض عند مستوى معنوية 5% أو 1% على أساس نتائج المعادلة (4.55).

وعلى سبيل توضيح إضافي، افترض أننا نرغب في اختبار الفرضية بأن b_1 ، أو MPC من دخل العمل المتاح، تساوي 0.5 مقابل الفرضية البديلة ذات الذيلين:

$$H_0: b_1 = 0.5,$$

$$H_1: b_1 \neq 0.5.$$

وب 22 درجة حرية ومستوى معنوية 5% نجد من الجدول الإحصائي رقم ٢ أن القيمة الحرجة لـ t هي 2.07، ومن المعادلة (4.55) تكون القيمة المحسوبة لنسبة t هي:

$$\frac{0.767 - 0.5}{0.047} = \frac{0.267}{0.047} = 5.7 > 2.07.$$

وعليه، سوف نرفض فرضية العدم عند مستوى معنوية 5% على أساس نتائج المعادلة (4.55)*. وفي الحقيقة، فإنه، عندما تكتسب بعض الخبرة في تحليل الانحدار، سوف تكتشف أن بإمكانك أداء كثير من هذه الاختبارات بوساطة الفحص الأولي. وعلى سبيل المثال، ففي الحالة السابقة، بلمحة واحدة لـ $b_1 (= 0.767)$ ، سوف يتضح لنا أنها تبعد عن القيمة المفترضة لـ b_1 (أو 0.5) بأكثر من ضعف الخطأ العشوائي (2×0.046). ونتيجة لذلك، فإن الفرضية بأن $b_1 = 0.5$ يمكن رفضها بدون إجراء أي حسابات.

ونلاحظ، أخيراً، أن المعادلة التي كونها أندو ومودجلياني لها قوة تفسيرية كبيرة. حيث إن $R^2 = 0.999$ ، وهذا يعني أنه يمكن أن ينسب للمتغيرين المستقلين 99% من التغير في القيم المشاهدة للاستهلاك. فإذا قبلنا الشكل العملي لهذه النظرية فإن هذه النتائج العملية تبدو متسقة مع نظريتهما للسلوك الاستهلاكي. ومن المهم أن ندرك، عند تفسير هذه النتائج، إمكانية أن تكون نتائج هذه المعادلة متسقة مع نظريات أخرى أيضاً. ذلك أنه، بينما يمكننا التصريح بأن النتائج العملية تتفق مع النظرية، فإنه لا يمكننا أن ندعي بأن نظريات الاستهلاك الأخرى غير صحيحة. ويبين هذا سبب كون التطبيق العملي على مشكلة معينة في الاقتصاد، عادة، عملية مستمرة، حيث تتراكم الأدلة التي تتفق مع سبب كون افتراض معين أو تتعارض معه. ذلك أن الدعم التطبيقي لفرضية ما يعتمد على المدى الذي تتسق فيه النتائج مع الفرضية. وبالدرجة نفسها من الأهمية على المدى الذي تكون فيه هذه النتائج غير متسقة مع الفرضيات المنافسة.

دراسة لضرائب المدينة

لختتم هذا الفصل، ستفحص دراسة أخرى عن الانحدار المتعدد، في هذه المرة باستخدام البيانات المقطعية. تظهر في هذه الأيام بعض المشاكل الحضرية

* ينبغي أن يكون واضحاً أن H_0 سوف ترفض، أيضاً، عند مستوى 1% من المعنوية.

المهمة، ويبحث رؤساء المدن عن مصادر للإيرادات الإضافية لتغطية نفقاتهم المتزايدة. واصبحت المشكلة تتلخص (إلى حد كبير) في محاولة فرض ضرائب جديدة أو رفع معدلات الضرائب القائمة دون إحداث خسارة مهمة في اقتصاد المدينة، خاصة أن هناك اقتناعاً لدى كثير من المراقبين أن المعدلات العالية من الضرائب قد عجلت من نزوح الطبقة المتوسطة إلى الضواحي، كما قللت شراء المستهلكين حاجياتهم من السلع والخدمات من المدن. أحد مصادر الإيرادات الرئيسة للمدن في الولايات المتحدة الأمريكية هي الضريبة على مبيعات التجزئة. يقال إن مثل هذه الضرائب تؤدي إلى انخفاض في مبيعات التجزئة وقد ساهمت في تدهور اقتصاديات المدن.

هل هذا صحيح، وإذا كان كذلك، هل له أهمية كمية كبيرة؟ وهذا، بالطبع، سؤال تصعب الإجابة عنه، ولكن دعنا نفكر بعض الشيء في كيفية استخدام التحليل القياسي لتناول هذه المشكلة، إذا أدت ضرائب المبيعات الأعلى في وسط المدن مقارنة بالضرائب في الضواحي إلى انخفاض المبيعات في المدينة. فسوف نتوقع أن نجد - مع بقاء الأشياء الأخرى على حالها - أنه، كلما كانت الاختلافات أكبر بين معدل الضرائب على المبيعات في مدينة معينة وبين الضرائب في ضواحيها، فسيزداد معدل المبيعات لكل نسمة في المدينة عن نظيرها في الضواحي. فإذا استطعنا الحصول على عينة من المدن تتوافر عنها بيانات عن مبيعات التجزئة لكل نسمة وحجم هذه الاختلافات الضريبية، فإنه قد يمكننا أن نكون انحدارا لمتغير المبيعات على متغير الاختلافات في المعدلات الضريبية، ثم نقوم، بفحص، بعد ذلك، العلاقة وحجم المعامل المقدّر والقيام بالاختبارات اللازمة.

ويبدو هذا المنهج معقولاً بدرجة كافية، إلا أن فيه مشكلة متأصلة تجعلنا لانشعر بالارتياح. ذلك أن فرضيتنا هي، مع بقاء الأشياء الأخرى على حالها، ارتباط اختلافات ضرائب المبيعات في المدن مع مستوى أقل من المبيعات. ولكن، في حدود عيتنا من المدن فإن الأشياء الأخرى لا تظل على حالها. ذلك أن مبيعات التجزئة في مدينة معينة تعتمد على عدد من المتغيرات المهمة إضافة إلى الضرائب،

كمستوى الدخل والحجم النسبي لسكان الضواحي. على سبيل المثال، إذا كان قاطنوا المدينة من الأغنياء يتوقع أن تكون المبيعات لكل نسمة أعلى، إضافة إلى ذلك، كلما ازداد عدد المشترين المحتملين في المدينة فستكون مبيعاتها أكبر. ولذا، فإن ما ينبغي علينا عمله هو التحكم في تأثير هذه المحددات الأخرى لمبيعات المدينة حتى يمكننا عزل تأثير الاختلافات في ضريبة المبيعات أو فصلها.

ويستدعي هذا، بالطبع أن نستخدم تحليل الانحدار المتعدد، وهنا يمكننا، بالتحديد، أن نكون انحدارا لمتغير المبيعات على مجموعة من المتغيرات المستقلة التي ستشتمل ليس فقط على متغير الضريبة ولكن أيضا على محددات مهمة للمبيعات في المدينة. مثل هذه الدراسة نفذها جون ميكسل John Mikesell*، والذي اقترح النموذج:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + b_3 X_{3t} + b_4 X_{4t} + u_t, \quad (4.56)$$

حيث:

Y_t = نصيب الفرد من مبيعات التجزئة في المدينة t .

X_{1t} = الاختلافات في ضريبة المبيعات للمدينة t .

X_{2t} = الدخل لكل نسمة في المدينة t .

X_{3t} = نسبة سكان المدينة t إلى إجمالي السكان في المدينة وضواحيها.

X_{4t} = مساحة المدينة t (بالأميال المربعة).

وقد عرف متغير الاختلافات في ضريبة المبيعات على النحو:

$$X_{1t} = \frac{1+t_c}{1+t_s},$$

حيث إن t_c هو معدل الضريبة على المبيعات في المدينة t_s هو معدل الضريبة المتوسط على المبيعات في البلديات المحيطة بها. لذلك، فزيادة معدل ضريبة

*"Central Cities and Sales Tax Rate Differentials : The Border City Problem." *National Tax Journal* 23 (June 1970), pp. 206-213

المبيعات في المدينة t مع عدم تغير t_s سوف يزيد X_{1t} وسيخفض طبقا للمعادلة (4.56) المبيعات لكل نسمة، طالما أننا نفترض $b_1 < 0$.

استطاع ميكسل أن يجمع بيانات لعينة من 173 مدينة رئيسة وضواحيها في الولايات المتحدة الأمريكية، واستخدم هذه البيانات لتقدير المعادلة (4.56) باستخدام الانحدار المتعدد ووجد أن:

$$\hat{Y} = 4.5 - 7.44X_1 + 0.43X_2 - 0.11X_3 - 0.08X_4 \quad N=173 \quad (4.57)$$

(2.94) (0.10) (0.04) 0.02 $R^2=0.26$

حيث الأرقام داخل الأقواس هي الأخطاء المعيارية المقدرة المناظرة. ونلاحظ مباشرة أن معامل متغير الضريبة X_1 له العلاقة السالبة المتوقعة وأكبر من ضعف حجم الخطأ المعياري المناظر، ويعني ذلك أن نسبة t تزيد على ٢. يوحى هذا بوجود علاقة سالبة بين مبيعات التجزئة في المدينة والاختلافات الضريبية، بمعنى أنه عند مستوى معنوية 5% يمكننا رفض فرضية العدم بأن $b_1 = 0$ وقبول الفرضية البديلة $b_1 < 0$. نلاحظ، أيضا، أن تقدير b_1 يوحى حجم هذا التأثير يكون عادة كبيرا، فزيادة ١٪ في الاختلافات الضريبية على المبيعات بين المدينة وضواحيها ترتبط بانخفاض قدره ٧٪ في المتوسط، تقريبا، في مبيعات التجزئة في المدينة الرئيسية.*

وينبغي علينا أن ندرك أن هذه النتائج مقنعة إقناعا أكبر من النتائج التي تترتب على الانحدار البسيط لمبيعات المدينة على متغير الضريبة. ومع نتائج

* تمثل المعادلة (4.56) تبسيطاً لنتائج ميكسل. ففي الحقيقة يقترح ميكسل علاقة تأخذ شكل حاصل ضرب على النحو التالي:

$$Y_t = b_0 X_{1t}^{b_1} X_{2t}^{b_2} X_{3t}^{b_3} X_{4t}^{b_4} e^{u_t},$$

وبأخذ اللوغاريتمات يصبح لدينا:

$$\ln Y_t = \ln b_0 + b_1 \ln X_{1t} + b_2 \ln X_{2t} + b_3 \ln X_{3t} + b_4 \ln X_{4t} + u_t.$$

ويجدر ذكر أننا سنعالج العلاقات التي تأخذ شكل حاصل ضرب في الفصل القادم، ولكن، نلاحظ أنها تعميم مباشر لمعالجتنا في الفصل السابق للتحويلات اللوغاريتمية.

الانحدار المتعدد هذا، أخذ ميكسل في الحسبان بصراحة تأثير الدخل الفردي ونسبة سكان المدينة إلى سكان الضواحي ومساهمة المدينة على المبيعات. وبمعنى آخر يمكننا القول أن تأثير الضريبة على مبيعات التجزئة بالمدينة قد قيس مع أخذ تأثير المتغيرات الأخرى في الحسبان. وهكذا فإن نتائج ميكسل تتفق مع المقولة بأن الاختلافات في ضريبة المبيعات تسبب تخفيضات مهمة في مبيعات التجزئة في المدن الرئيسية.

ملحق أ (A)

خصائص المقدرات

يهدف هذا الملحق إلى أن يطور، بدقة أكثر، المقدرات وتبايناتها ومعاملات الانحدار، ثم إثبات أن هذه المقدرات غير متحيزة. يضاف إلى ذلك، أن المنهج الذي سنتبعه سيزودنا بمعرفة إضافية حول كيفية قيام تحليل الانحدار المتعدد بفصل تأثيرات مختلف المتغيرات المستقلة.

افترضنا في الفصل الرابع أن بعض المتغيرات المستقلة لنموذج الانحدار ترتبط، في الأقل، بالمتغيرات الأخرى كافة أو بعضها. كما ذكرنا في ذلك الفصل، سنحاول تفسير المتغير المستقل الأخير بدلالة المتغيرات المستقلة الأخرى، عن طريق نموذج الانحدار:

$$X_{kt} = c_0 + c_1 X_{1t} + \dots + c_{k-1} X_{(k-1)t} + v_{kt}. \quad (4A.1)$$

افترض أننا نقدر معلمات (4A.1) بطريقة المتغير المساعد. وللقيام بذلك سنضع $\Sigma(\hat{v}_{kt} X_{(k-1)t}) = 0$ و \dots و $\Sigma(\hat{v}_{kt} X_{1t}) = 0$ و $\Sigma \hat{v}_{kt} = 0$ للحصول على مجموعة المعادلات الطبيعية:

$$\begin{aligned} \sum X_{kt} &= n\hat{c}_0 + \hat{c}_1 \sum X_{1t} + \dots + \hat{c}_{k-1} \sum X_{(k-1)t} \\ \sum (X_{kt} X_{1t}) &= \hat{c}_0 \sum X_{1t} + \hat{c}_1 \sum X_{1t}^2 + \dots + \hat{c}_{k-1} \sum (X_{1t} X_{(k-1)t}) \\ &\vdots \\ \sum (X_{kt} X_{(k-1)t}) &= \hat{c}_0 \sum X_{(k-1)t} + \hat{c}_1 \sum (X_{1t} X_{(k-1)t}) + \dots + \hat{c}_{(k-1)} \sum X_{(k-1)t}^2 \end{aligned} \quad (4A.2)$$

لاحظ وجود معلمات غير معلومة عددها $k: \hat{c}_0, \hat{c}_1, \dots, \hat{c}_{k-1}$ ومعادلات عددها k في المعادلة (4A.2) لذلك - وبتحقق افتراضاتنا - يمكننا (عموما) أن نحل هذه المعادلات للحصول على تلك المعلمات. ويمكننا هذا من تكوين:

$$\hat{X}_{kt} = \hat{c}_0 + \hat{c}_1 X_{1t} + \dots + \hat{c}_{k-1} X_{(k-1)t}. \quad (4A.3)$$

وسيكون مقدر v_{kt} هو:

$$\hat{v}_{kt} = X_{kt} - \hat{X}_{kt}. \quad (4A.4)$$

وبإعادة ترتيب الحدود في المعادلة (4A.4) يصبح لدينا:

$$X_{kt} = \hat{X}_{kt} - \hat{v}_{kt}. \quad (4A.5)$$

لاحظ، بدقة، مافعلناه حتى الآن. قسمنا قيمة X_k إلى جزئين: الجزء الأول \hat{X}_{kt} ، وهو ذلك الجزء من X_k الذي يرتبط مباشرة بالمتغيرات المستقلة الأخرى، فهو يمثل التغير في X_k الذي يمكن أن ينسب إلى المتغيرات المستقلة الأخرى X 's. وبالمقابل، فإن \hat{v}_{kt} هو ذلك الجزء من X_k الذي لانستطيع نسبته إلى بوساطة المتغيرات المستقلة الأخرى أو تفسيره بوساطتها. لاحظ، مرة أخرى، أنه إذا كان $\hat{v}_{kt} = 0$ لجميع قيم t فسوف يكون لدينا ارتباط متعدد خطي تام [انظر المعادلة (4A.5) و (4A.3)] وحينئذ، لا يمكننا تقدير b_k . ولهذا السبب، نتوقع أن يؤدي \hat{v}_{kt} دورا مهما في تقدير b_k .

وبعد أخذ هذا في الحسبان، دعنا نسترجع مماورد في هذا الفصل أنه لدينا عدد $(k+1)$ من المعادلات، الطبيعية المحددة لـ \hat{b}_i ، وتأخذ إحداها الشكل:

$$\sum (Y_t X_{kt}) = \hat{b}_0 \sum X_{kt} + \hat{b}_1 \sum (X_{1t} X_{kt}) + \dots + \hat{b}_k \sum X_{kt}^2. \quad (4A.6)$$

نعرف من المعادلة (4A.5) أن X_{kt} يمكن التعبير عنها كمجموع كل من \hat{X}_{kt} و \hat{v}_{kt} ، كما نعلم، أيضا، أن $\sum \hat{v}_k = 0$ و $\sum \hat{v}_{kt} X_{ik} = 0$ حيث $k=1, 2, \dots$ و $i=1, 2, \dots$. ولذلك طالما أن \hat{X}_{kt} توليفة خطية من $X_{1t}, \dots, X_{2k}, \dots, X_{(k-1)t}$ [انظر المعادلة (4A.3)] وأن $\sum (\hat{X}_{kt} - \hat{v}_{kt}) = 0$ فإذا عوضنا عن X_{kt} في كل مكان في المعادلة (4A.6) بـ $(\hat{X}_{kt} + \hat{v}_{kt})$ وألغينا الحدود التي تساوي الصفر كافة، فإننا سنحصل على:

$$\sum (Y_t \hat{X}_{kt}) + \sum (Y_t \hat{v}_{kt}) = \quad (4A.7)$$

$$\hat{b}_0 \sum \hat{X}_{kt} + \hat{b}_1 \sum (X_{1t} \hat{X}_{kt}) + \dots + \hat{b}_k \sum \hat{X}_{kt}^2 + \hat{b}_k \sum \hat{v}_{kt}^2.$$

ويتساوى مجموع الحدود الـ (k+1) الأولى في الجانب الأيمن من المعادلة (4A.7)، ببساطة، مع $\sum (Y_t \hat{X}_{kt})$:

$$\sum (Y_t \hat{X}_{kt}) = \hat{b}_0 \sum \hat{X}_{kt} + \hat{b}_1 \sum (X_{1t} \hat{X}_{kt}) + \dots + \hat{b}_k \sum \hat{X}_{kt}^2. \quad (4A.8)$$

وحتى نرى ذلك نعوض:

$$Y_t = \hat{b}_0 + \hat{b}_1 X_{1t} + \dots + \hat{b}_k X_{kt} + \hat{u}_t$$

في المجموع الأيسر من المعادلة (4A.8) حيث نلاحظ أن:

$$\sum (\hat{u}_t \hat{X}_{kt}) = \hat{c}_0 \sum \hat{u}_t + \hat{c}_1 \sum (\hat{u}_t X_{1t}) + \dots + \hat{c}_{(k-1)} \sum (\hat{u}_t X_{(k-1)t}) = 0.$$

وسبب كل هذا هو أننا إذا عوضنا المعادلة (4A.8) في المعادلة (4A.7) نحصل على:

$$\sum (Y_t \hat{v}_{kt}) = \hat{b}_k \sum \hat{v}_{kt}^2, \quad (4A.9)$$

ثم نحل المعادلة (4A.9) للوصول إلى \hat{b}_k :

$$\hat{b}_k = \frac{\sum (Y_t \hat{v}_{kt})}{\sum \hat{v}_{kt}^2} \quad (4A.10)$$

وباختصار، تعتمد \hat{b}_k فقط على Y_t 's و \hat{v}_{kt} حيث تمثل \hat{v}_{kt} التغير المستقل في X_{kt} وتعميم

المعادلة (4A.10) عموماً هو:

$$\hat{b}_i = \frac{\sum (Y_t \hat{v}_{it})}{\sum \hat{v}_{it}^2} \quad (4A.11)$$

حيث \hat{v}_{it} هو الباقي من انحدار X_{it} على X_t 's الأخرى كافة. أي أن مقدر كل واحد من b_i يمكن التعبير عنه بدلالة قيم Y_t وقيم الحد \hat{v}_{it} الذي يمثل التغير «المستقل»، في المتغير المستقل المقابل، X_{it} .

مقدرات غير متحيزة

في المبحث السابق، أثبتنا أن:

$$\hat{b}_i = \frac{\sum (Y_i \hat{v}_{it})}{\sum \hat{v}_{it}^2} \quad (4A.11)$$

ولكي نثبت أن \hat{b}_i غير متحيزة، نعيد كتابة النموذج الأساسي في الفصل والموضح في معادلة (4.3) على النحو:

$$Y_i = b_0 + b_1 X_{i1} + \dots + b_i (\hat{X}_{it} + \hat{v}_{it}) + \dots + b_k X_{ik} + u_i, \quad (4A.12)$$

حيث عوضنا عن X_{it} بـ $(\hat{X}_{it} + \hat{v}_{it})$. بعد ذلك نضرب المعادلة (4A.12) في \hat{v}_{it} ونجمع على مدى n من المشاهدات، ثم نعوض عن $\sum (Y_i \hat{v}_{it})$ في المعادلة (4A.11) للحصول على:

$$\hat{b}_i = b_i + \frac{\sum (\hat{v}_{it} \hat{u}_i)}{\sum \hat{v}_{it}^2} \quad (4A.13)$$

عند حصولنا على المعادلة (4A.13) استخدمنا شروط المعادلة الطبيعية $\sum \hat{v}_{it} = 0$ و $\sum (\hat{v}_{it} X_{it}) = 0$ إذا كانت $i \neq j$ ولذا تكون $\sum (\hat{v}_{it} \hat{X}_{it}) = 0$. لاحظ أن \hat{v}_{it} لا يعتمد على قيم u_i ولكن يعتمد، فقط، على القيم المعطاة لـ X 's. وطالما نفترض استقلال u_i عن X 's (لأي قيم معطاة لـ X 's) فإن \hat{v}_{it} سوف يحصل عليها وسيصبح لدينا:

$$\begin{aligned} E(\hat{b}_i) &= E(b_i) + E \left[\frac{\sum (\hat{v}_{it} \hat{u}_i)}{\sum \hat{v}_{it}^2} \right] \\ &= b_i + \left(\frac{\hat{v}_{i1}}{\sum \hat{v}_{it}^2} \right) E(u_1) + \dots + \left(\frac{\hat{v}_{in}}{\sum \hat{v}_{it}^2} \right) E(u_n) \\ &= b_i. \end{aligned} \quad (4A.14)$$

وهكذا، فمقدرنا لـ b_i غير متحيز.

تباينات المقدرات

والآن، من السهل علينا أن نشق التباين الشرطي لـ \hat{b}_i من المعادلة (4A.13) وبالتحديد دعنا نعيد كتابة المعادلة (4A.13) بالتفصيل للحصول على:

$$\hat{b}_i = b_i + \left(\frac{\hat{v}_{i1}}{\sum \hat{v}_{ii}^2} \right) u_1 + \left(\frac{\hat{v}_{i2}}{\sum \hat{v}_{ii}^2} \right) u_2 + \dots + \left(\frac{\hat{v}_{in}}{\sum \hat{v}_{ii}^2} \right) u_n \quad (4A.15)$$

وبجعل $M_{it}^2 = \hat{v}_{it} / \sum \hat{v}_{ii}^2$ يصبح لدينا:

$$\hat{b}_i = b_i + M_{i1} u_1 + \dots + M_{in} u_n \quad (4A.16)$$

أي أن \hat{b}_i توليفة خطية من الأخطاء العشوائية. وباستخدام الصيغة الموجودة في الفصل الثاني لتباين المجموع الخطي من المتغيرات العشوائية غير المرتبطة ببعضها بعضا، يكون لدينا:

$$\text{var}(\hat{b}_i) = M_{i1}^2 \sigma_u^2 + M_{i2}^2 \sigma_u^2 + \dots + M_{in}^2 \sigma_u^2. \quad (4A.17)$$

دع $A = \sum \hat{v}_{ii}^2$. حيث يكون $M_{ii}^2 = \hat{v}_{ii}^2 / A^2$. وباستخدام هذا يمكننا إعادة كتابة المعادلة (4A.17) على النحو:

$$\begin{aligned} \text{var}(\hat{b}_i) &= \frac{\sigma_u^2}{A^2} = (\hat{v}_{i1}^2 + \hat{v}_{i2}^2 + \dots + \hat{v}_{in}^2) \\ &= \frac{\sigma_u^2 (\sum \hat{v}_{ii}^2)}{A^2} \\ &= \frac{\sigma_u^2}{A^2} A = \frac{\sigma_u^2}{A} \end{aligned} \quad (4A.18)$$

باختصار،

$$\text{var}(\hat{b}_i) = \frac{\sigma_u^2}{\sum \hat{v}_{ii}^2}, \quad \text{for } i = 1, \dots, k \quad (4A.19)$$

ملحق ب (B): العلاقة بين R^2 و \bar{R}^2

أشرنا في هذا الفصل إلى أن $R^2 \geq \bar{R}^2$. هنا نثبت هذه القاعدة، اعتبر نموذج الانحدار الخطي التالي:

$$Y_i = b_0 + b_1 X_{1i} + \dots + b_k X_{ki} + u_i, \quad (4B.1)$$

بحد ثابت وعدد k من المتغيرات المستقلة. لذلك، فإن عدد المعاملات هو $p = k + 1$. إذا اعتبرنا النموذج الذي يكون فيه $k \geq 1$ ، حيث $k \geq 2$ ، تكون $p \geq 2$. لمثل هذا النموذج، دع R^2 و \bar{R}^2 معامل التحديد ومعامل التحديد المعدل على الترتيب. حيث $\bar{R}^2 < R^2$ ، فإننا سوف نثبت أدناه مايلي:

$$\bar{R}^2 < R^2 \quad (4B.2)$$

إلا إذا كانت $R^2 = 1$ (عندها تكون $\bar{R}^2 = R^2 = 1$). وطالما أن R^2 تكون، عادة، أقل من الواحد الصحيح، فإن النتيجة في المعادلة (4B.2) تشير إلى أن \bar{R}^2 ستكون، عادة، أقل من R^2 .

للحصول على المعادلة (4B.2)، لاحظ من المعادلة (4.52) أن \bar{R}^2 يمكن التعبير عنها على النحو:

$$\bar{R}^2 = \left(1 - \frac{ESS}{TSS}\right) + \frac{ESS}{TSS} \cdot \frac{ESS/(n-p)}{TSS/(n-1)}. \quad (4B.3)$$

وبما أن $R^2 = 1 - ESS/TSS$ ، فإنه ينتج عن ذلك أن $ESS/TSS = 1 - R^2$. لاحظ، أيضا، أن:

$$\frac{ESS/(n-p)}{TSS/(n-1)} = \frac{ESS}{TSS} \left(\frac{n-p}{n-1} \right). \quad (4B.4)$$

وينتج عن المعادلة (4B.3) بعدئذ، أن \bar{R}^2 يمكن التعبير عنها على النحو:

$$\begin{aligned} \bar{R}^2 &= R^2 + \frac{ESS}{TSS} - \left(\frac{ESS}{TSS} \right) \left(\frac{n-1}{n-p} \right) \\ &= R^2 + \left(\frac{ESS}{TSS} \right) \left(1 - \frac{n-1}{n-p} \right) \end{aligned}$$

$$= R^2 + (1 - R^2) \left(1 - \frac{n-1}{n-p} \right) \quad (4B.5)$$

$$= R^2 + (1 - R^2) \left(\frac{n-p}{n-p} - \frac{n-1}{n-p} \right)$$

$$= R^2 + (1 - R^2) \left(\frac{1-p}{n-p} \right)$$

$$= R^2 - (1 - R^2) \left(\frac{p-1}{n-p} \right)$$

من الواضح الآن أنه، إذا كان $R^2 = 1$ فإن $\bar{R}^2 = R^2$. أما إذا كانت $R^2 < 1$ و $p \geq 2$ فإن $(1 - R^2) (p-1) / (n-p) > 0$. وحيث، ينتج عن السطر الأخير من المعادلة (4B.5) أن $\bar{R}^2 < R^2$.

أسئلة

١- اعتبر نموذج الانحدار التالي:

$$Y_t = a_0 + a_1 X_{1t} + a_2 X_{2t} + u_t$$

(أ) أذكر الافتراضات المعتادة لهذا النموذج.

(ب) أذكر المعادلات الطبيعية مشيراً إلى الافتراض الذي يناظر كلا منها.

(ج) افترض أن عيّنتنا تحتوي على $n=100$ ، $\sum X_1 = \sum X_2 = \sum X_1 X_2 = 0$ ، وأن

$\sum Y = 10$ ، $\sum (Y X_1) = 30$ ، $\sum (Y X_2) = 20$ ، $\sum X_1^2 = 35$ ، وأخيراً

$\sum X_2^2 = 3$. قدر a_0 ، a_1 و a_2 .

٢- اعتبر النموذج:

$$Y_t = a_0 + a_1 X_{1t} + a_2 X_{2t} + a_3 (X_{1t} - X_{2t}) + a_4 X_{1t} X_{2t} + \varepsilon_t$$

ماهي الملمات التي لا يمكن تقديرها في ظل افتراضاتنا المعتادة ؟ ولماذا ؟

٣- اعتبر نموذج الانحدار التالي:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t$$

حيث إن مشاهدات X_{1t} ، X_{2t} و Y_t هي :

X_{1t}	X_{2t}	Y_t
1	2	7
2	1	8
1	3	5
3	1	6
1	2	4

اكتب المعادلات الطبيعية .

٤- يقال إن العائلات عالية الدخل ومتوسطته تنزح من المدن وتعيش في الضواحي بسبب الضرائب العالية نسبيا ومعدلات الجرائم العالية، وتكاليف الاسكان العالية، وأيضا، بسبب أنها ترغب في مكان أوسع للإقامة. كون نموذج انحدار يمكن استخدامه لاختبار مثل هذه الفرضية، اشرح المزايا النسبية للبيانات المقطعية والسلاسل الزمنية لاختبار مثل هذه الفرضيات.

٥- افترض أن: $D_{1t} = a_0 + a_1 p_{1t} + a_2 p_{2t} + \dots + a_k p_{kt} + b \bar{p}_t + c y_t + u_{1t}$ حيث D_{1t} هو الطلب على السلعة 1، P_{1t} : ثمن السلعة 1، p_{2t}, \dots, p_{kt} هي أسعار (k-1) من السلع الأخرى و $\bar{P}_t = \sum_{i=1}^k (P_{it}) / k$ هو المستوى العام للأسعار، و Y_t هو الدخل. باختصار، يعتقد بعض الناس أن الطلب على السلعة 1 يعتمد على مستوى سعرها أسعار السلع الأخرى ومستوى الأسعار العام والدخل.

(أ) هل يمكن أن تثار أي مشاكل في تقدير هذا النموذج ؟
(ب) هل يمكن تقدير أي من معلمات النموذج السابق؟ ماهي تلك المعلمات؟ وضع.

٦- اعتبر النموذج: $Y_t = b_0 + b_1 X_t + b_2 X_t^2 + \varepsilon$

(أ) هل المعادلة السابقة تعاني من تعدد العلاقات الخطية التام.
(ب) افترض أن لدينا المشاهدات التالية:

Y	-1	-1	2	4
X	0	1	2	5

اكتب المعادلات الطبيعية .

طرق أخرى في تحليل الانحدار المتعدد

عمّما في الفصل السابق طريقة التقدير الأساسية لنموذج انحدار المتغيرين إلى حالة المتغيرات المستقلة المتعددة. في هذا الفصل، سندرس بعض الطرق الإضافية التي يمكن استخدامها في تحليل الانحدار المتعدد. وبالتحديد، سنوسع أولا تحليلنا السابق بمعالجة حالة المتغيرات المبطة في الانحدار المتعدد، وفي هذا المجال، سنقدم ثلاث طرق لتقدير الاشكال المختلفة للعلاقات المبطة. سنقدم ثانيا: مفهوم المتغيرات «الصورية» وهي تلك المتغيرات التي تمكنا من حساب بعض التأثيرات النوعية التي تدخل في علاقتنا الاقتصادية. على سبيل المثال، يمكننا استخدام المتغيرات الصورية من الأخذ في الحسبان تأثير عوامل مثل الجنس والدين على أنواع معينة من السلوك. هذه الطريقة تمكنا من أن ندخل في تحليلنا متغيرات لا يمكن قياسها قياسا كميا تقليديا. وأخيرا، سنعود إلى مسألة الشكل رقمالدالي وسنرى كيفية استخدام تحليل الانحدار المتعدد لتقدير أنواع مختلفة من العلاقات.

(١-٥) تقدير العلاقات المبطة

في الفصل الثالث، أوضحنا أن نموذج انحدار المتغيرين يمكن أن يستخدم لتقدير معادلات يعتمد فيها المتغير التابع على قيمة المتغير المستقل في فترة سابقة.

على سبيل المثال، تذكر أننا اعتبرنا الحال، التي يكون فيها الإنفاق الاستهلاكي في فترة معينة يعتمد على الدخل المتاح في الفترة الزمنية السابقة على النحو التالي:

$$C_t = a + bY_{d(t-1)} + u_t. \quad (5.1)$$

ولكن، ليس هناك سبب ضروري لاعتماد الاستهلاك الحالي على الدخل المتاح في الفترة الزمنية السابقة مباشرة، فقط، ذلك أن مستويات الدخل في فترات سابقة، بالإضافة إلى مستواه الحالي قد تؤثر على الإنفاق الاستهلاكي. فإذا كان ذلك صحيحا فسيكون لدينا علاقة من الشكل:

$$C_t = a + b_0Y + b_1Y_{d(t-1)} + \dots + b_kY_{d(t-k)} + u_t. \quad (5.2)$$

يطلق على هذا النوع من العلاقات فترات الإبطاء الموزعة distributed lag، ويعني هذا أن قيمة المتغير التابع في أي فترة زمنية معطاة تعتمد على مجموع مرجح بالأوزان للقيم الماضية للمتغير المستقل. ويمكننا أن نتخيل، في هذه الحال، أن المتغير التابع C_t سيتكيف ببطء للقيمة الحالية للمتغير المستقل Y_{dt} وذلك بسبب الدفع الذاتي المتراكم من القيم الماضية المبطأة للمتغير المستقل.

ومثال نموذج رياضي لمثل هذه العلاقة بين C_t و Y_{dt} هو نموذج تعتمد فيه قيم الإنفاق الاستهلاكي في الفترة t على الدخل المتوقع في الفترة $t+1$ ، ولكن، حيث يحدد الدخل المتوقع، بدوره، على أساس مجموع مرجح بالأوزان لمستويات الدخل في الفترات السابقة. افترض، مثلا، أن:

$$C_t = a + bY_{d(t+1)}^e + u_t, \quad (5.3)$$

حيث $Y_{d(t+1)}^e$ هو الدخل المتوقع في الفترة $(t+1)$. افترض، أيضا، أن الدخل المتوقع هو مجموع مرجح بالأوزان من الدخول الحالية والماضية على النحو:

$$Y_{d(t+1)}^e = \alpha_0Y_{dt} + \alpha_1Y_{d(t-1)} + \dots + \alpha_kY_{d(t-k)}. \quad (5.4)$$

وبالتعويض من المعادلة (5-4) في المعادلة (5.3) نحصل على:

$$C_t = a + b_0Y_{dt} + b_1Y_{d(t-1)} + \dots + b_kY_{d(t-k)} + u_t, \quad (5.5)$$

حيث إن $b_0 = b\alpha_0$ و $b_1 = b\alpha_1$ وهلم جرا. وهكذا، سوف تستجيب C_t بتباطؤ

لـ Y_{dt} بسبب أن Y_{dt} هو، فقط، أحد العوامل التي تحدد $Y_{d(t+1)}^e$. وعلى أي حال، يمكن، نظريا، تقدير المعادلة (5.2) مباشرة بطريقة الانحدار المتعدد. افترض أننا نسلم بأن الإنفاق الاستهلاكي الحالي يعتمد على الدخل الحالي وعلى الدخل في السنوات التسع الماضية، لذلك يكون لدينا المعادلة:

$$C_t = a + b_0 Y_{dt} + b_1 Y_{d(t-1)} + \dots + b_9 Y_{d(t-9)} + u_t. \quad (5.6)$$

لتقدير المعادلة (5.6)، نحصل على بيانات من النوع الموجود في الجدول رقم (١-٥)، حيث يمثل كل صف مشاهدة مشتركة، ويمكننا استخدام طرق الانحدار المتعدد للحصول على a ، b_0 ، b_1 ، ...، b_9 . أي أن افترض وجود بيانات من الشكل الموجود في الجدول رقم (١-٥) يمكننا من إعادة تسمية $Y_{d(t-1)}$ أو تعريفها بأنها X_{1t} ، $Y_{d(t-2)}$ بأنها X_{2t} ، ...، $Y_{d(t-9)}$ بأنها X_{9t} ثم تقدير دالة الاستهلاك (5.6) كما لو أنها معادلة انحدار عادية من الشكل:

$$C_t = a + b_0 Y_{dt} + b_1 X_{1t} + b_2 X_{2t} + \dots + b_9 X_{9t} + u_t. \quad (5.7)$$

وتظهر مشاهداتنا X_{1t} ، ...، X_{9t} في الجدول رقم (١-٥)*.

جدول رقم (١-٥)

C_t	Y_{dt}	$Y_{d(t-1)}$	$Y_{d(t-2)}$...	$Y_{d(t-9)}$
C_{1950}	$Y_{d(1950)}$	$Y_{d(1949)}$	$Y_{d(1948)}$...	$Y_{d(1941)}$
C_{1951}	$Y_{d(1951)}$	$Y_{d(1950)}$	$Y_{d(1949)}$...	$Y_{d(1942)}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
C_{1970}	$Y_{d(1970)}$	$Y_{d(1969)}$	$Y_{d(1968)}$...	$Y_{d(1961)}$

وعلى الرغم من أن هذه العلاقة تعد شكلا ملائما للتقدير في بعض الحالات، إلا أنها تسبب بعض المشاكل. تذكر من الفصل الثالث أنه عندما أدخلنا فترة إبطاء

* يجب على القارئ أن يقتنع نفسه، على سبيل المثال، بأن $X_{2(1951)} = Y_{d(1949)}$

واحدة في علاقتنا ذات المتغيرين فقدنا مشاهدة واحدة. ولكن، في حالتنا هذه، فالوضع أسوأ لأننا نفقد مشاهدة لكل قيمة مبطأة اضافية للدخل المتاح تدخل المعادلة (5.2). افترض، مثلاً، أنه يتوافر لدينا قيم مشاهدة للاستهلاك وللدخل المتاح لعشر سنوات من ١٩٦٠م وحتى ١٩٦٩م، وأنا قدرنا المعادلة (5.6) حيث يعتمد الاستهلاك على الدخل الحالي وعلى الدخل في السنوات التسع الماضية. في هذه الحال، سنفقد تسعا من المشاهدات العشر. أي أن السنة الوحيدة التي تتوافر بها مشاهدات لجميع المتغيرات التي تظهر في المعادلة (5.6) ستكون ١٩٦٩م.

$$C_{1969} = a + b_0 Y_{d(1969)} + b_1 Y_{d(1968)} + \dots + b_9 Y_{d(1960)} \quad (5.8)$$

وتتطلب علاقة الاستهلاك السابقة لعام ١٩٦٩م استخدام بيانات عن الدخل المتاح قبل ١٩٦٠م غير أنها ليست متاحة. ونتيجة لذلك، فلن تتوافر لدينا مشاهدات كاملة يمكن من خلالها تقدير معادلتنا للانحدار.*

أحد المشاكل الأخرى المرتبطة بالنماذج من النوع (5.6) تنشأ إذا كانت k كبيرة (يعني ذلك، عادة، أن $k \geq 5$)، ويعني ذلك وجود عدد كبير من الملاحظات التي ينبغي تقديرها. يضاف إلى ذلك أن هذه الملاحظات ستناظر متغيرات ترتبط بصورة قوية ببعضها البعض. وكما يتوقع القارئ فإن ذلك سيجعل من الصعب عزل تأثير مختلف المتغيرات المستقلة على المتغير التابع (سنيين ذلك في الفصل القادم). بمعنى آخر، ففي ظل هذه الافتراضات، ستكون تباينات مقدرات معلمات الانحدار كبيرة.

* ينبغي أن يكون القارئ قادراً، على إثبات أنه إذا كان لدينا مشاهدة واحدة مشتركة فقط، فستكون هناك معادلة طبيعية مستقلة واحدة، لأن جميع المعادلات الطبيعية ستأخذ شكل نسبة من المعادلة الطبيعية الأولى. فمثلاً، بالنسبة للمعادلات (5.7) و (5.8) يصبح معامل النسبية في المعادلة الطبيعية الثانية هو وفي المعادلة الطبيعية الثالثة هو وهلم جرا [تلميح للحل ستصبح المشاهدة الوحيدة في المعادلة (5.7) أو (5.6) هي عندما يكون $t=1$ ، حيث تشير الفترة الزمنية 1 إلى السنة ١٩٦٩م وسيحصل على المعادلات الطبيعية من الشروط:

$$\sum_{t=1}^1 \hat{u}_t = \hat{u}_1 = 0, \sum_{t=1}^1 (\hat{u}_t Y_{dt}) = \hat{u}_1 Y_{d1} = 0, \sum_{t=1}^1 (\hat{u}_t X_{1t}) = \hat{u}_1 X_{11} = 0, \dots, \sum_{t=1}^1 (\hat{u}_t X_{9t}) = \hat{u}_1 X_{91} = 0.$$

وهكذا فإن المعادلة الطبيعية الثانية هي Y_{d1} مضروبة في الأولى، والثالثة هي X_{11} مضروبة في الأولى وهلم جرا. تمثل هذه النتيجة حالة خاصة من نتيجة أكثر عمومية تصرح بأنه، ينبغي أن يتوافر لدينا عدد k على الأقل من المشاهدات المشتركة لكي نستطيع تقدير المعلمات.

ثمة مشكلتان أساسيتان مرتبطتان بتحليل فترات الإبطاء الموزعة، الأولى هي المشاهدات التي تفقد بسبب فترات الإبطاء، والثانية هي وجود عدد كبير من المعلومات التي ينبغي تقديرها تقديرا يعتمد عليه. ولمواجهة هذه المشاكل طور الاقتصاديون نماذج لتحليل فترات الإبطاء الموزعة والتي إما أن تقلل من عدد المشاهدات التي تفقد بسبب الإبطاء و/ أو تقلل من عدد المعلومات التي ينبغي تقديرها. نعرض الآن نوعين من هذه النماذج.

إبطاء كويك The Koyck Lag

النموذج الأول هو نموذج، إبطاء كويك*. يستخدم هذا النموذج افتراضا يرتبط بمعلومات العلاقة ذات فترات الإبطاء الموزعة، ويسمح بترجمة هذه العلاقة إلى شكل أبسط كثيرا. وينتج عن هذا النموذج فترات إبطاء أقل وعددا أقل من المعلومات التي يتم تقديرها. ولكن، لسوء الحظ، فإن هذا الشكل رقما لا بسط من العلاقة ينتج عنه تعقيدات خطيرة كثيرا ما يتم تجاهلها. ولأن نموذج كويك قد شاع استخدامه، فإننا نعرض له هنا ونبين أوجه قصوره، يضاف إلى ذلك أن عرضه سيخدمنا باعتباره خطوة مهمة في المناقشات التالية.

افترض أنه، على الرغم من أن الاستهلاك يعتمد على الدخل في السنوات السابقة فإن تأثير الدخل في السنوات الماضية البعيدة أقل من تأثير الدخل في السنوات الأخيرة. وبالتحديد، افترض أن الاستهلاك الحالي هو مجموع مرجح بالأوزان من السنوات الحالية والماضية للدخل المتاح (مضافا إليه الخطأ العشوائي) وأن هذه الأوزان تتناقصا متتاليا للفترات الأكثر بعدا في الماضي. يفترض نموذج كويك تناقص هذه الأوزان هندسيا. على سبيل المثال، اجعل λ ثابتا تتراوح قيمته بين الصفر والواحد الصحيح. حيث λ ، وبالرجوع إلى المعادلة (5.2)، يأخذ نموذج كويك الشكل:

* يرجع إلى: L. M. Koyck, *Distributed Lags and Investment Analysis* (Amsterdam: North Holland, 1954).

$$b_i = \lambda^i b_0, \quad i = 1, 2, \dots, k. \quad (5.9)$$

وبالتعويض من المعادلة (5.9) في المعادلة (5.2)، يصبح لدينا:

$$C_t = a + b_0 Y_{dt} + (b_0 \lambda) Y_{d(t-1)} + (b_0 \lambda^2) Y_{d(t-2)} + \dots + b_0 \lambda^k Y_{d(t-k)} + u_t. \quad (5.10)$$

تبين المعادلة (5.10) أن الاستهلاك يعتمد على كل من مستويات الدخل الحالية والماضية، ولكن طالما أن λ مرفوعة إلى قوى أعلى فإنها تصغر باستمرار، وهكذا، تصغر معاملات السنوات الأقدم بالتوالي كلما توغلنا في الماضي.

سيوضح لنا الآن متضمنات نموذج الإبطاء لكويك. فإذا قمنا بإبطاء المعادلة (5.10) لفترة زمنية واحدة، ثم ضربنا المعادلة الناتجة في λ نحصل على:

$$\begin{aligned} \lambda C_{t-1} &= \lambda a + (\lambda b_0) Y_{d(t-1)} \\ &+ (\lambda^2 b_0) Y_{d(t-2)} + \dots + (\lambda^{k+1} b_0) Y_{d(t-k-1)} + \lambda u_{t-1}. \end{aligned} \quad (5.11)$$

وبطرح المعادلة (5.11) من المعادلة (5.10)، نحصل على:

$$C_t - \lambda C_{t-1} = (a - \lambda a) + b_0 Y_{dt} - \lambda^{k+1} b_0 Y_{d(t-k-1)} + (u_t - \lambda u_{t-1}) \quad (5.12)$$

وبإعادة ترتيب جدول (5.12) يصبح لدينا:

$$C_t = (a - \lambda a) + b_0 Y_{dt} + \lambda C_{t-1} - (\lambda^{k+1} b_0) Y_{d(t-k-1)} + (u_t - \lambda u_{t-1}). \quad (5.13)$$

وبفرض أن k كبيرة (بمعنى وجود عدد كبير من سنوات الإبطاء في النموذج)، فسيكون الحد قبل الأخير في المعادلة (5.13) $(\lambda^{k+1} b_0) Y_{d(t-k-1)}$ صغيراً. ومن ثم، وعلى سبيل التقريب، سنضع هذا الحد مساوياً للصفر*. لتبسيط الرموز أيضاً، دع $a^* = (a - \lambda a)$ حيث يمكن تبسيط المعادلة (5.13) إلى:

$$C_t = a^* + b_0 Y_{dt} + \lambda C_{t-1} + (u_t - \lambda u_{t-1}). \quad (5.14)$$

والآن دع:

$$v_t = (u_t - \lambda u_{t-1}). \quad (5.15)$$

* وبالطبع إذا افترضنا أن k تؤول إلى مالانهاية بينما يظل الحد $Y_{d(t-k-1)}$ محدوداً فإن هذا الحد سيساوي الصفر. هذا هو الافتراض المعتاد الذي يستخدم في الدراسات القياسية.

طالما أن v_t يعتمد، فقط، على الأخطاء العشوائية، فمن المنطقي اعتبار v_t نفسه خطأ عشوائيا، دعنا للحظة، نفترض أن v_t (والذي، لسوء الحظ، غالبا ما يفترض في الدراسات القياسية) يحقق افتراضات نموذج الانحدار المرتبطة بخصائص الخطأ العشوائي كافة. ويمكن التعبير في هذه الحال، عن المعادلة (5.14) على النحو التالي:

$$C_t = a^* + b_0 Y_{dt} + \lambda C_{t-1} + v_t, \quad (5.16)$$

حيث a ثابت وأن:

$$E(v_t) = 0,$$

$$E(v_t^2) = \sigma_v^2,$$

$$E(v_t v_{t-1}) = 0,$$

$$E(v_t Y_{dt}) = E(v_t C_{t-1}) = 0.$$

لاحظ حجم التبسيط الهائل الذي تمثله المعادلة (5.16) مقارنة بالمعادلة (5.2) حيث جمع تأثير القيم المبطة كافة للدخل المتاح على الاستهلاك الحالي في حد واحد: قيمة الاستهلاك المبطة لفترة زمنية واحدة. نحتاج الآن، فقط، لتقدير قيمة λ بدلا من معاملات القيم المبطة كافة، وبمعنى آخر إذا قبلنا نموذج كويك لفترات الإبطاء، وقبلنا، أيضا، الافتراضات المرتبطة بـ v_t ، فإن معلمات نموذج مثل (5.2) يمكن تقديرها بدلالة نموذج مثل (5.16)، والذي يتطلب تقديره خسارة مشاهدة واحدة فقط.

وحتى نكون أكثر تحديدا لطريقة التقدير، يمكن اعتبار المعادلة (5.16) نموذجا للانحدار المتعدد يتضمن المعلمات a^* ، b ، وأخيرا λ . وسنحصل على المقدرات \hat{a}^* ، \hat{b}_0 ، وأخيرا $\hat{\lambda}$ وذلك بالتطبيق المباشر لطريقة المتغير المساعد. وباستخدام هذه المعلمات المقدرة نحصل على مقدرات، المعلمات كافة في المعادلة (5.2). فعلى سبيل المثال، إذا افترضنا أنه يوجد عدد لانهائي من فترات الإبطاء (أي k لانهائية):

$$\hat{b}_i = (\hat{\lambda}^i) \hat{b}_0, \quad i = 1, 2, \dots, \quad (5.17)$$

وحيث إن $a^* = a - \lambda a$ ، ولذلك فإن مقدر a سيكون:

$$\hat{a} = \frac{\hat{a}^*}{1 - \hat{\lambda}} \quad (5.18)$$

وللتعبير عن النتائج بعمومية أكثر، تمكنا افتراضات نموذج كويك (مشملة على افتراض أن k لانهائية) من وضع النموذج:

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + \dots + u_t \quad (5.19)$$

في الشكل المبسط:

$$Y_t = a^* + b_0 X_t + \lambda Y_{t-1} + v_t, \quad (5.20)$$

حيث نحتاج، فقط، لتقدير ثلاث معالم a^* ، b_0 ، و λ . وتصبح العلاقة بين معالم المعادلة (5.19) والمعادلة (5.20) هي:

$$a = \frac{a^*}{1-\lambda}, \quad b_i = \lambda^i b_0, \quad i=1,2,\dots \quad (5.21)$$

وهكذا، فعن طريق تحويل النموذج صراحة إلى نموذج يحتوي على فترة إبطاء، واحدة، يؤدي استخدام نموذج كويك لفقد مشاهدة مشتركة واحدة. ومزية نموذج كويك ينبغي أن تكون الآن واضحة.

ولكن - وكما لاحظنا من قبل - هناك بعض المشاكل الصعبة مع نموذج كويك. تذكر أولا أننا عند اشتقاق المعادلة التي تقدرها، افترضنا أن الخطأ العشوائي

$$v_t = (u_t - \lambda u_{t-1})$$

يحقق الشروط كافة التي تفرضها عادة الأخطاء العشوائية. ولكن، لسوء الحظ، فإن ذلك ليس صحيحا بالضرورة. فإذا كانت u_t في المعادلة الأصلية (5.19) تحقق افتراضات نموذج الانحدار كافة، فإن v_t في المعادلة (5.20) لا تحققها، عموماً. وبالتحديد، فإن قيم v_t سيكون لها ارتباط غير صفري مع بعضها بعضاً. وأيضاً، مع واحد من المتغيرات المستقلة - القيمة المبطة للمتغير التابع Y_t . على سبيل المثال، إذا كانت $v_t = u_t - \lambda u_{t-1}$ و $v_{t-1} = u_{t-1} - \lambda u_{t-2}$ ، فإن v_t و v_{t-1} لن يكونا مستقلين عن بعضهما بعضاً طالما أنهما يحتويان على حد مشترك (u_{t-1}). ومن ثم، لانتوقع أن تتحقق $E(v_t v_{t-1}) = 0$. وفي الحقيقة، فإنه، في ظل افتراضاتنا العادية المتعلقة بـ u_t ، لدينا:

$$\begin{aligned}
E(v_t v_{t-1}) &= E[(u_t - \lambda u_{t-1})(u_{t-1} - \lambda u_{t-2})] \\
&= E(u_t u_{t-1} - u_t \lambda u_{t-2} - \lambda u_{t-1}^2 + \lambda^2 u_{t-1} u_{t-2}) \\
&= -\lambda \sigma_u^2 \neq 0.
\end{aligned} \tag{5.22}$$

وهكذا فإن التباين بين القيم المتعاقبة للأخطاء العشوائية في المعادلة (5.20) لن يساوي الصفر. ونترك للقارئ، على سبيل التدريب، أن يثبت بطريقة مشابهة أن*:

$$E(v_t Y_{t-1}) \neq 0. \tag{5.23}$$

وباختصار، إذا بدأنا بمعادلة من الشكل رقم (5.19) والتي تحقق افتراضات نموذج الانحدار كافة فإن تحويل كويك لهذه المعادلة سيؤدي، عموماً إلى انتهاك بعض هذه الافتراضات.** يضاف إلى ذلك أن النتائج المترتبة على هذه الانتهاكات خطيرة. فمثلاً - وكما سنوضح في فصل لاحق - تتضمن المعادلة (5.23) أن مقدرات b_i و λ لن تكون متحيزة وحسب بل غير متسقة أيضاً.***

بالإضافة إلى مشاكل التقدير التي سبق الإشارة إليها، يعد نموذج كويك مفيداً جداً من حيث إنه يفترض تناقص تأثير الفترات الماضية تأثيراً متعاقباً وبطريقة معينة. وهذا، بالتأكيد، لن يكون صحيحاً دائماً. فعلى سبيل المثال، - وبسبب تأثير العادة - فقد يكون صحيحاً أن مستوى الدخل في الفترة السابقة مباشرة أكثر أهمية في التأثير على مستوى الاستهلاك الحالي عن مستوى الدخل الحالي. ومثل هذه العلاقة لا تتناسب بوضوح مع نموذج كويك لفترات الإبطاء. وهكذا يصبح من المفيد جداً الحصول على نموذج لفترات الإبطاء الموزعة يتصف بمرونة أكبر من نموذج كويك، كما لا يخالف افتراضات نموذج الانحدار.

* تلميح للحل: بما أن Y_t من المعادلة (5.20) تعتمد مباشرة على v_t ، فإن Y_{t-1} تعتمد على v_{t-1} و v_t و Y_{t-1} غير مستقلين عن بعضهما بعضاً. لإثبات المعادلة (5.23)، اضرب Y_{t-1} في v_t ثم خذ التوقعات. وعند القيام بذلك لاحظ أن $E[Y_{t-2} v_t] = 0$.

** سنعرض الطرق التي تستخدم لمواجهة انتهاكات افتراضات نموذج الانحدار في الفصلين السادس والسابع. *** لأن المعادلة الطبيعية الثالثة من نموذج الانحدار (5.20) مبنية على الشرط $\sum (\hat{v}_t Y_{t-1}) = 0$ ، الذي لم يعد متسقاً مع المعادلة (5.23). سنعرض مزيداً من التوضيح فيما بعد.

إبطاء آلمون *The Almon Lag

على الرغم من أن المشاكل التي تظهر في نموذج كويك يمكن مواجهتها في إطار نموذج أكثر عمومية سنكونه فيما بعد، فإن الحل ليس سهلاً. ويمكن التغلب على هذه الصعوبة عن طريق استخدام نموذج المون. ونؤكد هنا أن هذا النموذج لا يقلل - كما هو الحال في طريقة كويك - عدد المشاهدات التي تفقد بسبب وجود المتغيرات المبطأة، ولكنه يقلل، فعلاً، عدد المعلمات التي تقدر. إضافة إلى ذلك، هناك ميزتان لنموذج المون على نموذج كويك. الأولى أنه لا يخالف أي من افتراضات نموذج الانحدار. والثانية - وكما سنرى - أنه أكثر مرونة بدرجة كبيرة من نموذج كويك بدلالة أشكال الإبطاء وهياكله المسموح بها.

بالعودة إلى التكوين العام للعلاقة المبطأة، دعنا نفترض أن النموذج الذي نرغب في تقريره يأخذ - مرة أخرى - الشكل:

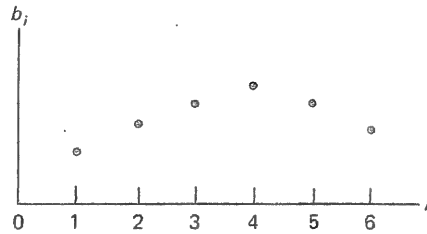
$$Y_t = a + b_0 X_t + b_1 X_{t-1} + \dots + b_k X_{t-k} + u_t, \quad (5.24)$$

حيث يستوفي الخطأ العشوائي u_t الافتراضات المعتادة. وكما سبق، قد نتوقع أن تكون معاملات الـ X 's المناظرة لفترات زمنية بعيدة أقل من تلك المعاملات المناظرة للفترات الزمنية الأقرب. ولكن، من ناحية أخرى، ولا اعتبارات معينة، فإن ذلك لن يكون صحيحاً بالضرورة في نموذج (5.24)، ففي بعض الحالات تزيد b 's فعلاً في البداية (أي $b_1 > b_0$) ثم تبدأ بعد ذلك في التناقص تدريجياً. فمثلاً، بسبب البطء في الإدراك، يحدث تباطؤ زمني في جمع البيانات، أو في عنصر الزمن المتضمن في اتخاذ القرارات، فالإنفاق الاستثماري، مثلاً، قد يكون أكثر استجابة لظروف الطلب في فترات زمنية قليلة سابقة عن ظروف الطلب الحالية. وباختصار، وفي ظل افتراضات مختلفة، يمكن أن نتوقع أنماطاً مختلفة لقيم b 's في نموذج مثل (5.24).

* انظر:

Shirley Almon. "The Distributed Lag between Capital Appropriations and Expenditures." *Econometrica*, 33 (Jan. 1965), pp. 178-196, and Ray Fair and Dwight Jaffee. "A Note on the Estimation of Polynomial Distributed Lags", *Econometrica Research Memorandum*, No. 120, Princeton University, Feb. 1971.

وعلى العكس من طريقة كويك، لا تفترض طريقة آلمون مثل هذه العلاقات الجامدة بين b 's. بدلا من ذلك تفترض أنه أيا كان نمط الـ b 's فإن ذلك النمط يمكن تحديده بمتعدد الحدود (polynomial). وعلى سبيل المثال، إذا توقعنا أن b 's تزداد في البداية ثم تتناقص بعد ذلك، فإن هذا النمط قد يشبه ذلك الموجود في الشكل رقم (١-٥).



شكل رقم (١-٥)

افترض أنه يمكننا تكوين منحنى متصل يمر بالنقاط في الشكل رقم (١-٥) كما فعلنا في الشكل رقم (٢-٥). نرغب الآن في وصف المنحنى بالشكل رقم (٥-٢) جبريا. في الرياضيات توجد نظرية تصرّح بأنه-وفي ظل شروط عامة- يمكن تقريب منحنى بوساطة متعدد الحدود. والقاعدة لتحديد درجة متعدد الحدود* هي أن تكون تلك الدرجة المختارة أكبر بمقدار واحد، في الأقل، من عدد نقاط الانقلاب Inflection Points في المنحنى. وبتطبيق هذه القاعدة على الشكل رقم (٢-٥) يمكننا تقريب المنحنى بوساطة متعدد الحدود من الدرجة الثانية، وبالتحديد وباستخدام طريقة آلمون، سنفترض:

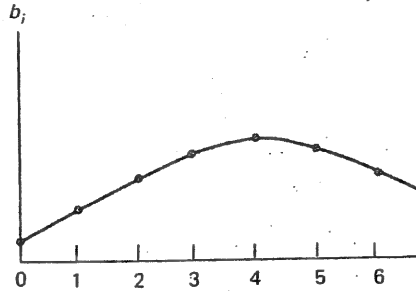
* تذكر من علم الجبر أن «درجة متعدد الحدود» تشير إلى أعلى قوة يرفع إليها المتغير. وهكذا فإن:

$$Y = b_0 + b_1X + b_2X^2$$

هو متعدد الحدود من الدرجة الثانية بينما:

$$Y = b_0 + b_1X + b_2X^2 + b_3X^3$$

هو متعدد الحدود من الدرجة الثالثة.



شكل رقم (٥-٢)

$$b_i = a_0 + a_1 i + a_2 i^2, \quad (5.25)$$

حيث إن α_0 ، α_1 و α_2 ثوابت ينبغي تحديدها. لاحظ أنه، إذا كانت المعادلة (5.25) هي تقريب للمنحنى في الشكل رقم (٥-٢) فإنه ينبغي أن يكون لدينا:

$$\begin{aligned} b_0 &= \alpha_0 & (set \ i=0) \\ b_1 &= \alpha_0 + \alpha_1 + \alpha_2 & (set \ i=1) \\ b_2 &= \alpha_0 + 2\alpha_1 + 2\alpha_2 & (set \ i=2) \\ &\vdots \\ b_k &= \alpha_0 + k\alpha_1 + k^2\alpha_2 & (set \ i=k) \end{aligned} \quad (5.26)$$

وتشتق صيغة كل من b_i في المعادلة (5.26) مباشرة وببساطة من المعادلة (5.25) عن طريق وضع i مساوية قيمة الدليل السفلي لمعامل معين. قد تبدو المعادلة (5.25) غريبة بعض الشيء عند النظر إليها من حيث إن قيمة الوزن المبثثة b_i يرتبط بطول الفجوة الزمنية ذاتها i . وفي الحقيقة، فإننا واجهنا علاقة مشابهة في طريقة كويك. فهناك افترضنا أن:

$$b_i = \lambda^i b_0. \quad (5.27)$$

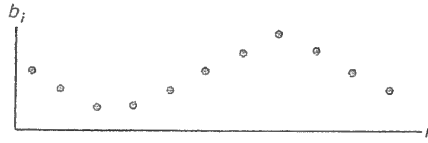
في هذه المعادلة، يرتبط b_i مرة أخرى بـ i . والاختلاف الوحيد بين المعادلة (5.27) والمعادلة (5.25) هو شكل المعادلة.

قبل أن ننفذ طريقة ألون، دعنا نوضح، باختصار، إلى أي مدى تتسم هذه الطريقة بالمرونة. افترض أننا نشعر بأن b 's تتبع نمطا مثل ذلك الموجود في الشكل رقم (٣-٥) حيثند سنفترض، ببساطة، أن:

$$b_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \alpha_3 i^3, \quad (5.28)$$

التي تتضمن أن:

$$\begin{aligned} b_0 &= \alpha_0, \\ b_1 &= \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3, \\ b_2 &= \alpha_0 + 2\alpha_1 + 4\alpha_2 + 8\alpha_3, \\ &\vdots \\ b_k &= \alpha_0 + k\alpha_1 + k^2\alpha_2 + k^3\alpha_3. \end{aligned} \quad (5.29)$$



شكل رقم (٣-٥)

وبشمولية أكثر، لاستخدام طريقة ألون، كل ما علينا فعله هو حساب عدد نقاط الانقلاب في نمطنا المفترض لـ b 's، ثم نعبر عن b 's بوصفه متعدد الحدود في i ، حيث تكون درجة متعدد الحدود واحد مضاف إليها عدد نقاط الانقلاب في المنحنى. في الشكل رقم (٤-٥) رسمنا عددا من أنماط فترات الإبطاء الممكنة مع درجة متعدد الحدود المناظر.

سنبين الآن كيفية استخدام طريقة آلمون لتقدير العلاقة التي تنطوي على فترة إبطاء. دعنا الآن نعود إلى التكوين العام الذي اعتبرناه سابقا:

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + \dots + b_k X_{t-k} + u_t, \quad (5.24)$$

افترض أن النظرية الاقتصادية توحي أن متعدد الحدود من الدرجة الثانية مناسب لتحديد شكل فترة الإبطاء. سنأخذ في هذه الحالة:

$$b_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2. \quad (5.30)$$

فإذا عوضنا عن b's في المعادلة (5.24) بوساطة صيغها الموجودة في المعادلة (5.30) نحصل على:

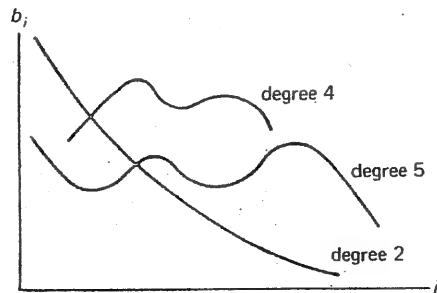
$$Y_t = a + \alpha_0 X_t + (\alpha_0 + \alpha_1 + \alpha_2) X_{t-1} + (\alpha_0 + 2\alpha_1 + 4\alpha_2) X_{t-2} + \dots + (\alpha_0 + k\alpha_1 + k^2\alpha_2) X_{t-k} + u_t. \quad (5.31)$$

وبإعادة ترتيب الحدود في المعادلة (5.31)، نحصل على:

$$Y_t = a + \alpha_0 \left(\sum_{i=0}^k X_{t-i} \right) + \alpha_1 \left(\sum_{i=1}^k i X_{t-i} \right) + \alpha_2 \left(\sum_{i=1}^k i^2 X_{t-i} \right) + u_t. \quad (5.32)$$

دعنا الآن نبسط رموزنا عن طريق تعريف:

$$Z_{1t} = \sum_{i=0}^k X_{t-i}, \quad Z_{2t} = \sum_{i=1}^k i X_{t-i}, \quad \text{and} \quad Z_{3t} = \sum_{i=1}^k i^2 X_{t-i}. \quad (5.33)$$



شكل (٤-٥)

حيث يمكننا إعادة كتابة المعادلة (5.32) على النحو:

$$Y_i = a + \alpha_0 Z_{1i} + \alpha_1 Z_{2i} + \alpha_2 Z_{3i} + u_i. \quad (5.34)$$

المعادلة (5.34) هي نموذج انحدار متعدد عادي يربط Y_i بـ Z_{1i} ، Z_{2i} و Z_{3i} . وبسهولة، يمكن أن نوجد مقدرات لكل من a ، α_0 ، α_1 و α_2 بطريقة التقدير العادية. دع \hat{a} ، $\hat{\alpha}_0$ ، $\hat{\alpha}_1$ و $\hat{\alpha}_2$ هي هذه المقدرات. حيثئذ، يمكننا أن نرى من المعادلة (5.30) أن مقدراتنا لـ b 's هي:

$$\begin{aligned} \hat{b}_0 &= \hat{\alpha}_0, \\ \hat{b}_1 &= \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2 \\ \hat{b}_2 &= \hat{\alpha}_0 + 2\hat{\alpha}_1 + 4\hat{\alpha}_2 \\ &\vdots \\ \hat{b}_k &= \hat{\alpha}_0 + k\hat{\alpha}_1 + k^2\hat{\alpha}_2. \end{aligned} \quad (5.35)$$

لاحظ أنه، بهذه الطريقة، أمكننا الحصول على مقدرات لعدد k من الملمات b_0, \dots, b_k عن طريق الحصول على مقدرات لهذه الملمات الثلاثة α_0 ، α_1 وأخيرا α_2 . والآن يمكن إثبات أنه، في أي حالة يكون فيها عدد أكبر من b 's عن الـ α 's تكون مقدرات b 's التي حُصل عليها من طريقة آلمون أفضل (بمعنى أنها تحتوي على تباينات أصغر) من المقدرات المباشرة لـ b 's التي حُصل عليها بوساطة تطبيق طريقة الانحدار المتعدد مباشرة على المعادلة (5.24).^{*} ولسوء الحظ فإننا لانستطيع اعطاء صيغة مبسطة لتباين المقدرات b_0, \dots, b_k التي يحُصل عليها بوساطة طريقة آلمون. ولكن في التطبيق سيزودنا الحاسوب بتقديرات لهذه التباينات حتى نقوم بالاختبارات الاحصائية المعتادة المرتبطة بقيم معاملات الانحدار.

إن تعميم طريقة آلمون واستخدام صيغ مختلفة منها سهل ومباشر. افترض،

^{*} كما نتوقع، فإن هذه النتيجة تكون مبنية على افتراض أن العلاقة المفترضة بين الـ b 's والـ α 's مثل المعادلة (5.30) صحيحة.

مثلا، أن المعادلة (5.24) قد وسعت لتشتمل على متغير آخر:

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + \dots + b_k X_{t-k} + cW_t + u_t, \quad (5.36)$$

حيث W_t هو متغير مستقل آخر. إذا افترضنا مرة أخرى أن b 's تحدد بعلاقة مثل (5.30) فإنه يمكننا أن نتبع الخطوات نفسها لاختزال المعادلة (5.36) لمعادلة تماثل المعادلة (5.24) مع استثناء وحيد وهو وجود الحد cW_t :

$$Y_t = a + \alpha_0 Z_{1t} + \alpha_1 Z_{2t} + \alpha_2 Z_{3t} + cW_t + u_t. \quad (5.37)$$

أي أن ادخال المتغيرات الإضافية لا يؤثر أي تأثير على تحليلنا. وفي الحقيقة، يمكننا أن نطبق طريقة آلمون لكل من المتغيرات المستقلة المبثثة العديدة في المعادلة نفسها. ينبغي علينا تذكر ملاحظتين إضافيتين مرتبطتين باستخدام منهج الإبطاء لآلمون، الأولى هي أن المستخدم قد يرغب في وضع بعض القيود الطرفية endpoints على قيم b 's فقد نرغب في جعل أما أن b_0 أو b_k (أو كليهما) تعادل الصفر. وبسبب التأخير في تلقي المعلومات، فقد نعتقد، مثلا، أن قيمة المتغير المستقل في الفترة الجارية لا تؤثر على السلوك الحالي (أي على قيمة المتغير التابع في هذه الفترة)، أي أنه في المعادلة (5.34)، نجعل $b_0 = 0$. ويعني هذا أن المتغير التابع (Y) سيعتمد على القيم المبثثة لـ X في المعادلة (5.24). ومن الناحية الأخرى، وربما بسبب الطريقة التي تتخذ بها القرارات، قد نعتقد أن قيم X المبثثة k أو أكثر من الفترات لا تأثير لها على Y ، لذلك، نجعل $b_k = 0$ في المعادلة (5.24).

إن إحدى طرق تضمين النموذج مثل $b_k = 0$ هي، ببساطة، إسقاط المتغير X_{t-k} من المعادلة الأساسية (5.24)، وإكمال التحليل كما سبق. ولكن، ليست هذه هي الطريقة التي تتبع، عادة. فعلى الصعيد العملي، تترجم المعلومات التي تبين أن $b_0 = 0$ أو $b_k = 0$ أو كليهما، باستخدام الافتراض الأساسي مثل المعادلة (5.30) شرطا لـ α 's، ثم تقدير المعادلة الناتجة حيثئذ. وعلى الرغم من أن إثبات ذلك يتجاوز مجال هذا الكتاب، فإن هذا المنهج غير المباشر يستخدم لأنه في ظل تحقق بعض الافتراضات يصبح تبين المقدرات غير المتحيزة الناتج أصغر من نظيره إذا اتبع المنهج المباشر المتلخص في إسقاط X_{t-k} و X_t . وعلى القارئ المهتم أن ينظر إلى الملحق أ (A) لهذا

الفصل حيث تعالج هذه المسألة بتفصيل أكبر، ويوضح الملحق كيف يمكن لمنهج آلمون تضمين هذه الشروط، ولكنه يقترح، أيضا، أسبابا لإتباع المنهج المباشر في التطبيق. ولهذا بعض الأهمية لأن معظم برامج الحاسوب لآلمون تتطلب من المستخدم تحديد القيود الطرفية في حال وجودها.

الثانية: لا بد أنك قد لاحظت أننا قد عرضنا لمادة هذا الجزء كما لو أن الباحث يعرف كلا من فترة الإبطاء في نموذج الانحدار (أي قيمة k) والنمط العام لـ b 's وذلك لتحديد درجة متعدد الحدود. إلا أنه، في التطبيق، قد لانعرف أيا من k أو نمط الـ b 's. في مثل هذه الحال، نصح باتباع المنهج التالي: أولا، حدد درجة لمتعدد للحدود (مثلا d) عالية بدرجة كافية حتي تلائم أي نمط معقول لـ b 's. في معظم الحالات الدرجة الثالثة أو الرابعة لمتعدد الحدود تكون كافية. افترض أن فترة الإبطاء القصوى «المعقولة» التي نعتقد باتساقها مع العلاقة موضع الاهتمام هي k^* . وعلى سبيل المثال، إذا تم استخدام بيانات ربع سنوية، فإن k^* قد تكون 12 أو 16 والتي تناظر 3 إلى 4 سنوات فترات إبطاء. أما إذا استخدمت بيانات شهرية فربما نأخذ k^* لتعادل 36. وعلى أي حال فعند اختيارك لـ d و k^* قدر العلاقة موضع الاهتمام عند $(k = d, d + 1, \dots, k^*)$ حيث d هي درجة متعدد الحدود. لاحظ أننا نهتم فقط بفترات الإبطاء حيث $k \geq d$ لأننا نفترض أن طول فترة الإبطاء تكون على الأقل بطول d (هناك على الأقل عدد من الـ b 's بقدر عدد الـ α 's).^{*} ينبغي أن تقدر كافة معادلات الانحدار التي تناظر القيم المختلفة لـ k بالبيانات نفسها. لاحظ أن هذا يتطلب أن نهمل المشاهدات الـ k^* الأولى وأن نستخدم فقط الـ $(n - k^*)$ المتبقية من المشاهدات لتقدير معادلات الانحدار. وسيسمح لنا هذا بمقارنة احصائية R^2 لمختلف المعادلات طالما أنها تؤسس على العينة نفسها. وهكذا ينبغي أن تأخذ قيمة k التي تعظم إحصائية R^2 . وفي ظل افتراضات معينة يمكن إثبات أن هذا المنهج يؤدي إلى مقدرات متسقة لكل من k ومعاملات الانحدار.

^{*} تذكر أن هدف طريقة آلمون للإبطاء هو تقليل عدد العلامات التي ينبغي تقديرها. وهذا لن يحدث إذا كانت $k < d$.

مثال

لشرح طريقة آلمون في التقدير ببساطة نعود إلى دالتنا السابقة للاستهلاك. في الجدول رقم (٥-٢) نعيد كتابة المشاهدات العشر حول الإنفاق الاستهلاكي والدخل المتاح للسنوات ١٩٦٠-١٩٦٩م والتي استخدمناها لتقدير معادلتنا التوضيحية للاستهلاك في الفصل الثاني (انظر الجدول رقم ٢-٢). سنستخدم هذه البيانات مرة أخرى لأغراض التقدير. ولكن، في حالتنا هذه، نفترض أن الإنفاق الاستهلاكي يعتمد على الدخل المتاح ذا فترات الإبطاء الموزعة. وأكثر تحديدا دعنا نفترض أن الاستهلاك يعتمد على الدخل المتاح في السنة الحالية وفي السنوات الأربع السابقة. إضافة إلى ذلك، نفترض أن فترة الإبطاء تأخذ شكل متعدد الحدود من الدرجة الثانية. وهكذا تأخذ دالتنا للاستهلاك الشكل رقم التالي:

$$C_t = a + b_0 Y_{dt} + b_1 Y_{d(t-1)} + b_2 Y_{d(t-2)} + b_3 Y_{d(t-3)} + b_4 Y_{d(t-4)} + u_t, \quad (5.38)$$

حيث:

$$b_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2.$$

جدول رقم (٥-٢) الاستهلاك والدخل المتاح في الولايات المتحدة بـ ١٠٠ مليون دولار الجارية

السنة	الاستهلاك (C)	الدخل المتاح Y_d
١٩٦٠	٣٢٥	٣٥٠
١٩٦١	٣٣٥	٣٦٤
١٩٦٢	٣٥٥	٣٨٥
١٩٦٣	٣٧٥	٤٠٥
١٩٦٤	٤٠١	٤٣٨
١٩٦٥	٤٣٣	٤٧٣
١٩٦٦	٤٦٦	٥١٢
١٩٦٧	٤٩٢	٥٤٧
١٩٦٨	٥٣٧	٥٩٠
١٩٦٩	٥٧٦	٦٣٠

المصدر: التقرير الاقتصادي للرئيس، واشنطن، D.C، مكتب الطباعة الحكومي، فبراير ١٩٧٠، الصفحات ١٨٩ و ١٩٥.

* تذكر أن هدف طريقة آلمون للإبطاء هو تقليل عدد العلامات التي ينبغي تقديرها. وهذا لن يحدث إذا كانت $k < d$.

ولوضع معادلتنا للاستهلاك في شكل المون ينبغي أن نحسب:

$$Z_{1t} = \sum_{i=0}^4 Y_{d(t-i)}, \quad Z_{2t} = \sum_{i=1}^4 i Y_{d(t-i)}, \quad \text{and} \quad Z_{3t} = \sum_{i=1}^4 i^2 Y_{d(t-i)}.$$

وتظهر قيم Z 's مع قيم المتغير التابع في الجدول رقم (٣-٥). لاحظ أننا نتيجة وجود فترة إبطاء لأربع سنوات نفقد أربعاً من مشاهداتنا، ولذلك فالجدول رقم (٣-٥) يحتوي على بيانات عن ست سنوات فقط. ولتوضيح الحسابات، نحصل على القيمة لـ $Z_{3(1969)}$ عن طريق حساب:

$$\begin{aligned} Z_{3(1969)} &= Y_{d(1968)} + 4Y_{d(1967)} + 9Y_{d(1966)} + 16Y_{d(1965)} \\ &= 590 + 2,188 + 7,568 = 14,954. \end{aligned}$$

وباستخدام البيانات الجديدة في الجدول رقم (٣-٥) يمكننا استخدام طريقتنا العادية لتقدير المعادلة:

$$C_t = a + \alpha_0 Z_{1t} + \alpha_1 Z_{2t} + \alpha_2 Z_{3t} + u_t. \quad (5.39)$$

وتكون المعادلة المقدرة هي:

$$\hat{C}_t = -43.5 + 1.02Z_{1t} - 1.44Z_{2t} + 0.35Z_{3t}. \quad (5.40)$$

وبمساعدة قيمنا المقدرة لـ α_i ، يمكننا حساب تقديرات b_i :

$$\begin{aligned} \hat{b}_0 &= \hat{\alpha}_0 = 1.02, \\ \hat{b}_1 &= \hat{\alpha}_0 + \hat{\alpha}_1 = 1.02 - 1.44 = -0.07, \\ \hat{b}_2 &= \hat{\alpha}_0 + 2\hat{\alpha}_1 + 4\hat{\alpha}_2 = 1.02 + 2(-1.44) + 4(0.35) = -0.46, \\ \hat{b}_3 &= \hat{\alpha}_0 + 3\hat{\alpha}_1 + 9\hat{\alpha}_2 = 1.02 + 3(-1.44) + 9(0.35) = -0.15, \\ \hat{b}_4 &= \hat{\alpha}_0 + 4\hat{\alpha}_1 + 16\hat{\alpha}_2 = 1.02 + 4(-1.44) + 16(0.35) = 0.86. \end{aligned}$$

وهكذا تكون معادلتنا المقدرة للاستهلاك ذات فترات إبطاء المون الأربع

هي: *

* في هذه الحال، لا تتوافق المعادلة المقدرة توافقا جيدة مع توقعاتنا. وتوحي العلامات الجبرية لمعاملات القيم البطأة لمتغير الدخل ضرورة التفكير جيداً عند صياغة دالة الاستهلاك.

$$\hat{C}_t = -43.5 + 1.02Y_{dt} - 0.07Y_{d(t-1)} + 0.46Y_{d(t-2)} \quad (28)$$

$$-0.15Y_{d(t-3)} + 0.86Y_{d(t-4)} \quad R^2 = 0.99 \quad (5.41)$$

جدول رقم (٣-٥)

Year	Ct	Z _{1t}	Z _{2t}	Z _{3t}
1964	401	1,942	3,667	10,821
1965	433	3,065	3,859	11,347
1966	466	2,213	4,104	12,030
1967	492	2,375	4,392	12,826
1968	537	2,560	4,742	13,860
1969	576	2,752	5,112	14,954

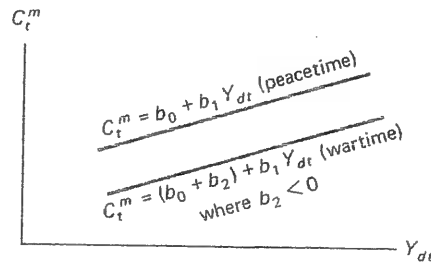
إضافة إلى القيم المقدرة للمعاملات، فقد عرضنا، أيضاً، القيم المطلقة لنسب t ، ولعامل التحديد. وتظهر معظم برامج الحاسوب التي تتضمن استخدام منهج آلمون للتقدير هذه المعلومات الإضافية.

(٢-٥) استخدام المتغيرات الصورية

حتى الآن، تعاملنا تعاملًا مكثفًا مع المتغيرات التي يمكن قياسها في وحدات كمية، مثل مستوى الدخل المتاح، أو معدل التغير في معدلات الأجور. ولكن في كثير من الأحيان، نعتقد بوجود متغيرات معينة ذات أهمية عظيمة ولكن، من طبيعة نوعية. فقد نعتقد، على سبيل المثال، أن مستوى الإنفاق الاستهلاكي الإجمالي يعتمد ليس فقط على مستوى الدخل المتاح، ولكن، أيضاً، على ما إذا كان المجتمع في حالة حرب أم في حالة سلم. ولكن كيف ندخل متغيراً، لوقت السلم أو وقت الحرب في معادلة الانحدار.*

* لمناقشة أكثر تعمقاً من المناقشة التالية، انظر:

إحدى طرق مواجهة هذه المشكلة هي تقدير معادلتين منفصلتين للاستهلاك، وذلك عن طريق استخدام بيانات فترة الحرب لتقدير دالة الاستهلاك في وقت الحرب، واستخدام بيانات وقت السلم لتقدير دالة الاستهلاك في وقت السلم، وهكذا فسنحصل على معادلتين مختلفتين للاستهلاك. ولكن، هناك طريقة أكثر كفاءة وهي تقدير معادلة واحدة للفترتين ولكن بعد وضع بعض الافتراضات. دعنا نفترض أن القيود على الاستهلاك وقت الحرب لا تغير الميل الحدي للاستهلاك، ولكنها تخفض الميل المتوسط له. ووفقا للشكل (٥-٥) نفترض أن دالة الاستهلاك خلال سنوات الحرب لها الميل نفسه كما في سنوات السلم، ولكن لها قاطع أدنى (أو حد ثابت أصغر). باستخدام هذا الافتراض نستطيع التعبير عن دوال الاستهلاك وقت الحرب ووقت السلم بدلالة معادلة انحدار واحدة هي:



الشكل رقم (٥-٥)

$$C_t = b_0 + b_1 Y_{dt} + b_2 D_t + u_t, \quad t = 1, \dots, n, \quad (5.42)$$

حيث:

 $D_t = 0$ لسنوات السلم $D_t = 1$ لسنوات الحرب

تبين المعادلة (5.42) أنه خلال سنوات السلم عندما تكون $D_t = 0$ يصبح لدينا:

$$C_t = b_0 + b_1 Y_{dt} + u_t, \quad (5.43)$$

بينما في فترات الحرب، عندما تكون $D_t = 1$ فإنه يصبح لدينا:

$$C_t = (b_0 + b_2) + b_1 Y_{dt} + u_t, \quad (5.44)$$

حيث يفترض أن $b_2 < 0$.

افترض أن الفترة الزمنية موضع الاهتمام هي حيث تكون فترة سنوات الحرب من $t=5$ إلى $t=9$. في هذه الحال، وفقا للمعادلة (5.42)، لدينا مجموعة من البيانات مثل الموجودة في الجدول رقم (٥-٤) وباستخدام هذه البيانات يمكننا تقدير قيم المعاملات في المعادلة (5.42) بطريقة الانحدار المتعدد العادية.

افترض أننا فعلنا ذلك وحصلنا على المعادلة:

$$\hat{C}_t = 40 + 0.9 Y_{dt} - 30 D_t, \quad (5.45)$$

جدول رقم (٥-٤)

t	الدخل المتاح	الإنفاق الاستهلاكي	D
1	Y_{d1}	C_1	0
.	.	.	.
.	.	.	.
4	Y_{d4}	C_4	0
5	Y_{d5}	C_5	1
5	Y_{d6}	C_6	1
7	Y_{d7}	C_7	1
8	Y_{d8}	C_8	1
9	Y_{d9}	C_9	1
10	Y_{d10}	C_{10}	1
.	.	.	.
.	.	.	.
.	.	.	.
n	Y_{dn}	C_n	0

حيث يتضح من نسبة t المناظرة إلى المتغير D_t أنها ذات حجم كاف مما يوحي بأن المعلمة b_2 في المعادلة (5.42) غير صفيرية. ومن ثم يجب أن نستنتج بأن الحرب لها تأثير سالب ومعنوي على الإنفاق الاستهلاكي. وستصبح معادلة الاستهلاك المقدرة لسنوات السلم وسنوات الحرب هي على الترتيب.

$$\hat{C}_t = 40 + 0.9Y_{dt} \quad (5.46)$$

و

$$\hat{C}_t = 10 + 0.9Y_{dt}, \quad (5.47)$$

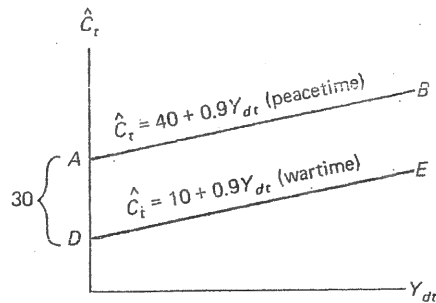
فإذا قيست النفقات الاستهلاكية ببلاتين الدولارات، فإن المقارنة بين المعادلة (5.46) و المعادلة (5.47) توضح أنه، عند مستويات الدخل المختلفة، تنخفض النفقات الاستهلاكية بمقدار 30 بليون دولار خلال سنوات الحرب. ويوضح الشكل رقم (٥-٦) هذه الدوال، حيث نرى أن دالة الاستهلاك لوقت الحرب DE هي خط مستقيم بميل دالة الاستهلاك نفسه لوقت السلم AB ولكن مع قاطع رأسي يقل بمقدار ٣٠ عن AB. وبالمقابل، قد نفترض أن ظروف وقت الحرب تقلل الميل الحدي للاستهلاك دون الحد الثابت في معادلة الاستهلاك.* في هذه الحال، تأخذ معادلتنا للانحدار المشتملة على كلتا الفترتين الشكل:

$$C_t = b_0 + b_1 Y_{dt} + b_2 (Y_{dt} D_t) + u_t, \quad (5.48)$$

حيث لدينا مرة أخرى

$D_t = 0$ لسنوات السلم،

$D_t = 1$ لسنوات الحرب.



شكل (٥-٦)

* في التطبيق، يمكن للفرد أن يقرر ما إذا كان القاطع أو بالمقابل، الميل الحدي للاستهلاك هو الذي ينتقل عن طريق دراسة أشكال أنواع القيود المفروضة على الاستهلاك وقت الحرب ... إلخ.

تبين المعادلة (5.48) بأن دالة الاستهلاك في أوقات السلم هي:

$$C_t = b_0 + b_1 Y_{dt} + u_t, \quad (5.49)$$

وذلك طالما أن $D_t = 0$ ، بينما تكون في أوقات الحرب

$$C_t = b_0 + (b_1 + b_2) Y_{dt} + u_t, \quad (5.50)$$

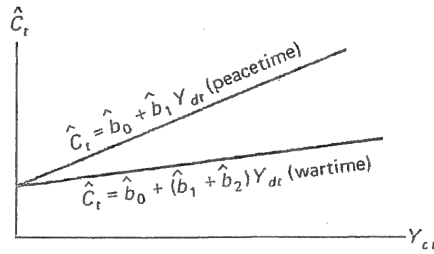
ونتوقع أن $b_2 < 0$. وكما سبق، فإنه يمكن استخدام بيانات مثل تلك الموجودة في الجدول رقم (٥-٤) لتقدير المعادلة (5.48). وستشابه العلاقة المقدرة الناتجة تلك المنحنيات الموجودة في الشكل رقم (٥-٧)، حيث يكون لدالة الاستهلاك في وقت الحرب انحدار أقل، ولكن القاطع الرأسي نفسه كما هو الحال في وقت السلم.

ويطلق على المتغير D الذي يظهر في المعادلة أعلاه «المتغير الصوري» dummy variable، وهو يأخذ قيمة الواحد الصحيح إذا تحققت بعض الشروط وقيمة الصفر إذا لم تحقق. وفي الحقيقة فإن استخدام المتغيرات الصورية تعميم قوي جدا لتحليل الانحدار، وكما سنري فإنه يسمح لنا تعميم مجال تحليلنا للانحدار ليشمل متغيرات مهمة لا يمكننا وصفها في وحدات كمية. وهكذا فباستخدام المتغيرات الصورية يمكن الأخذ في الحسبان تأثير العوامل النوعية المهمة التي تؤثر على متغيرنا التابع.

اعتبر، مرة أخرى، المثال الأول أعلاه، حيث افترضنا غير القاطع وثبات الميل الحدي للاستهلاك (م ح س) خلال سنوات الحرب، وبدلاً من استخدام طريقة المتغيرات الصورية، قمنا بعمل دالتين للاستهلاك، واحدة لسنوات الحرب وأخرى لسنوات السلم. في هذه الحال، سنقدر أربع معلمات وليس ثلاثاً. (في ظل افتراضنا بأن (م ح س) هو نفسه لوقت الحرب ووقت السلم فسوف ننتهي بتقديرين لمعلمة واحدة. وتصبح مشكلتنا هي استخدام هذين التقديرين للحصول على تقدير وحيد «أفضل» لـ (م ح س). في مثل هذه الحال، سنحاول، على الأرجح، تطوير بعض الطرق للحصول على القيمة المتوسطة لهذين التقديرين.

يمكن إثبات أنه إذا دمج هذان المتغيران معاً بطريقة مثلى (يجب تحديدها) فإن النتيجة تكون متماثلة مع التقدير الذي يتحقق بوساطة طريقة المتغير الصوري تحققت أكثر دقة، b_{1p} هو مقدار الـ (م ح س) المبني على معادلة وقت السلم وبياناته، دع،

b_{1w} أيضا، المقدّر المناظر المشتق من معادلة وقت الحرب وبياناته. في ظل تحقيق افتراضاتنا العادية، فإن هذه المقدرات غير متحيزة. فإذا دمجت هذه المقدرات دمجا ينجم عنه مقدّر غير متحيز لـ (م ح س) مع أصغر التباينات الممكنة، فإن المقدّر الناتج يكون متماثلا مع المقدّر الذي يتحقق بوساطة المتغير الصوري. وفي الحقيقة، فإن طريقة المتغير الصوري تستخدم معلومات العينة المتاحة كافة - إضافة إلى المعلومات المسبقة المرتبطة بتغيرات المعلمة - بأفضل الطرق الممكنة.



شكل (٧-٥)

ويمكننا - بهذه المناسبة - استخدام عدد كبير من المتغيرات الصورية بقدر ما نريد، شرط أن يتوافر لدينا عدد كاف من المشاهدات يسمح بتقدير المعادلة. افترض، على سبيل المثال، أننا نرغب في تفسير السلوك الاستهلاكي للأسر مختلفة. وأننا نعتقد أن مستوى استهلاك هذه الأسر يعتمد على عدد من السمات المميزة لها، كوجود الأطفال أو غيابهم، وما إذا كانت الأسرة تقيم في منزل تملكه أو تستأجره. وعنصر أرباب الأسر، وهلم جرا، إضافة إلى الدخل المتاح. فإذا استطعنا الحصول على هذه المعلومات كافة لعينة من الأسر، فإنه يمكن، على سبيل المثال، تقدير المعادلة:

$$C_t = b_0 + b_1 Y_{dt} + b_2 F + b_3 H_t + b_4 R_t + b_5 A_t + u_t, \quad (5.51)$$

حيث:

$$C_t = \text{الإنفاق الاستهلاكي للأسرة } t.$$

$$Y_{dt} = \text{الدخل المتاح للأسرة } t.$$

$$F_t = 1 \text{ إذا كانت الأسرة لها أطفال.}$$

$$0 \text{ إذا كانت الأسرة بدون أطفال.}$$

$$H_t = 1 \text{ إذا كانت الأسرة تملك المنزل الذي تقيم فيه.}$$

$$0 \text{ إذا كانت الأسرة لا تملكه.}$$

$$R_t = 1 \text{ إذا كانت الأسرة من العنصر الأبيض}$$

$$0 \text{ إذا كانت الأسرة غير ذلك}$$

$$A_t = 1 \text{ إذا كانت رب الأسرة يتجاوز الخمسين عاما}$$

$$0 \text{ إذا كان رب الأسرة غير ذلك.}$$

$$u_t = \text{الخطأ العشوائي.}$$

مثال

إن تطبيقات المتغيرات الصورية، في الحقيقة غير محدودة. دعنا نأخذ مثالا آخر يتضمن نوعا مختلفا تماما من المشاكل، والمثال من دراسة حديثة لأحد مؤلفي هذا الكتاب.* والقضية موضوع الدراسة هي ما إذا كانت الطبيعة الرسمية للدستور السياسي للدولة لها تأثير منتظم على درجة اللامركزية في المالية العامة للدولة. أو باختصار، هل الدستور عامل مهم في تحديد النصيب النسبي للنشاط المالي للحكومة المركزية في القطاع العام ككل؟

بعد أن تحققنا من أهمية متغيرات أخرى كحجم السكان، ومستوى الدخل لكل نسمة، فإن الطريقة هي إدخال متغير صوري بقيمة الواحد الصحيح إذا كانت

* أنظر: W. E. Oates. *Fiscal Federalism* (New York: Harcourt Brace Jovanovich, 1972), Chap. 5.

للدولة دستور فيدرالي (أي يضمن بعض الاستقلال الذاتي للمستويات الحكومية المحلية) أو قيمة الصفر في غياب الدستور الفيدرالي (حيث يحدد مجال السلطات الحكومية المحلية بوساطة الحكومة المركزية). باستخدام البيانات المقطعية لعينة من ٥٣ دولة، كانت المعادلة المقدرة هي:

$$\hat{G} = 96 - 1.21 \ln P - 0.004Y - 0.6Z - 15.9F \quad N = 53, \quad R^2 = 0.65 \quad (5.52)$$

(12.1) (1.3) (2.3) (5.5) (4.7)

(حيث إن الأرقام الموجودة داخل الأقواس تحت المعاملات المقدرة هي القيم المطلقة لنسبة t)، وإن:

G = نصيب الحكومة المركزية من الإيرادات العامة الكلية (بالنسبة المئوية)،
 $\ln P$ = اللوغاريتم الطبيعي لعدد السكان (بالآلاف)،
 Y = متوسط دخل الفرد بالدولارات الأمريكية ١٩٦٥ م.
 Z = مساهمات الضمان الاجتماعي بوصفها نسبة مئوية من إجمالي الإيرادات العامة الجارية،

F = 1 للدول ذات الدساتير الفيدرالية،

0 للدول ذات الدساتير غير الفيدرالية.

تتفق نتائج المعادلة (5.52) بوضوح مع الافتراض بأن وجود دستور فيدرالي يسهم في زيادة درجة اللامركزية في المالية العامة. ويظهر معامل المتغير الصوري F بإشارة سالبة، كما أن نسبة t المناظرة له تزيد على أربعة. لذا، يمكننا أن نرفض، بسهولة، فرضية العدم القائلة بعدم وجود ارتباط بين G و F عند مستوى المعنوية 5%. كما يدل حجم المعامل بعد الأخذ في الحسبان تأثير حجم السكان، الدخل، وغيرها، على أن الحكومة المركزية في الدول الفيدرالية تحصل في المتوسط على 16% أقل من إجمالي الإيرادات العامة التي تحصل عليها الحكومات المركزية في الدول غير الفيدرالية بنسبة 16%. وهكذا تؤدي الفيدرالية المفروضة بوساطة الدساتير السياسية دورا مهما في تحديد درجة اللامركزية في النشاط التمويلي العام.

بعض النتائج الإضافية

بينت الدراسات أن المتغيرات الصورية مفيدة جدا في فصل الاختلافات الفصلية والإقليمية في السلوك. فمبيعات السيارات نتيجة لإدخال نماذج جديدة في فصل الأعياد، أو حجم الإنتاج من المحاصيل المختلفة التي تعتمد على ظروف الطقس، من الواضح أنها تتغير تغيرا منتظما مع فصول السنة، فإذا تعاملنا مع بيانات ربع سنوية أو شهرية فإنه يمكننا ادخال متغيرات صورية تناظر مختلف الفصول لتأخذ في الحسبان تأثيراتها. وبالمثل، عندما تتوقع وجود اختلافات إقليمية في السلوك، فإنه يمكننا أن نسمح بذلك عن طريق إدخال المتغيرات الصورية لمختلف الأقاليم. ولكي نعرف كيف يتم ذلك (ولنشير إلى أحد المحاذير التي ينبغي تجنبها) اعتبر المثال التالي: افترض أننا نحاول فهم العلاقة بين حجم الإنتاج من إحدى السلع وحجم العمل المطلوب لإنتاجها حيث نعتقد بوجود قوى فصلية منتظمة تؤثر في هذه العلاقة. لذا، قد يمكننا وضع النموذج على النحو:

$$Q_t = b_0 + b_1 L_t + b_2 S_t + b_3 H_t + b_4 F_t + b_5 W_t + u_t, \quad (5.53)$$

حيث:

Q_t = وحدات الإنتاج في الفصل (ربع السنة) t .

L_t = وحدات عنصر العمل

S_t = 1 للفصل السنوي ابريل - يونيو.

0 للفصول الأخرى

H_t = 1 للفصل السنوي يوليو - سبتمبر

0 للفصول الأخرى

F_t = 1 للفصل السنوي اكتوبر - ديسمبر

0 للفصول الأخرى.

W_t = 1 للفصل السنوي يناير - مارس

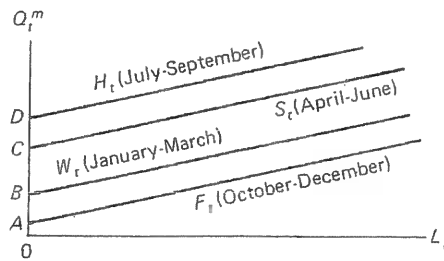
0 للفصول الأخرى.

لاحظ أننا أدخلنا متغيراً عشوائياً لكل واحد من الفصول ربع السنوية الأربعة

في السنة الميلادية. وبدلالة الشكل رقم (٥-٨) يبين نموذجنا أن القيمة المتوسطة لمستوى الإنتاج، أو Q_t^m ، تعادل $(b_0 + b_1 L_t)$ حيث إن b_1 هو الميل المشترك للخطوط الأربعة مضافا إليه كمية إضافية (يمكن أن تكون سالبة) تعتمد على الفصل ربع السنوي.

افترض أننا حاولنا تقدير المعادلة (5.53)، سوف نجد أنها، في شكلها العام، لا يمكن تقديرها بسبب أنها تتضمن ارتباطا متعددًا خطيًا تامًا بين المتغيرات المستقلة. أي أن افتراضنا بأن المتغيرات المستقلة ليست مرتبطة فيما بينها تماما وخطية قد خولف وبالتحديد فإن لدينا:

$$S_t + H_t + A_t + W_t \equiv 1 \quad (5.54)$$



شكل (٥-٨)

نعلم أنه أن أحد هذه المتغيرات، في أي فصل ربع سنوي معين، يأخذ قيمة الواحد الصحيح بينما البقية تأخذ قيمة الصفر، فمجموعهما ينبغي أن يكون الواحد دائما. تذكر من الفصل السابق أنه في وجود الارتباط الخطي المتعدد التام فلن نستطيع إيجاد تقديرات وحيدة للمعاملات لأنه لا يتوافر لدينا عدد كاف من المعادلات الطبيعية. ولكن هذه المشكلة يمكن حلها بسهولة عن طريق إسقاط أحد المتغيرات الصورية من المعادلة، وتغيير بعض تفسيراتها. افترض، على سبيل المثال، أننا نسقط W_t من المعادلة (5.53) للحصول على:

$$Q_t = b_0 + b_1 L_t + b_2 S_t + b_3 H_t + b_4 F_t + u_t, \quad (5.55)$$

والآن، فقد أزلنا الاعتماد الخطي بين المتغيرات المستقلة، وسنستطيع تقدير المعلمات

الخمس في المعادلة (5.55) خلال شهور الشتاء (يناير-مارس)، وستكون S_t ، H_t و F_t مساوية الصفر، كما سيكون الحد الثابت لمعادلتنا هو b_0 أي أن b_0 ستناظر القاطع الرأسي OB في الشكل رقم (٨-٥). وعلى نحو مشابه، وبالرجوع إلى المعادلة (5.55)، نجد أن المعامل لكل متغير صوري يشير إلى كيفية اختلاف تأثير الفصل المناظر عن تأثير فصل الشتاء. على سبيل المثال، فخلال فصل الربيع (أبريل - يونيو)، عندما يأخذ المتغير S_t قيمة الواحد الصحيح، فإنه يكون لدينا:

$$Q_t = (b_0 + b_2) + b_1 L_t + u_t \quad (5.56)$$

وفي الشكل رقم (٨-٥) يساوي القاطع الرأسي للمنحنى والمناظر للفصل (أبريل - يونيو) -OC- ($b_0 + b_2$) وعليه فإن b_2 يشير لاتجاه اختلاف أثر فصل الربيع عن فصل الشتاء وحجمه. فمثلا في الشكل رقم (٨-٥) يتوقع أن تكون b_2 موجبة وبالمثل من الشكل ذاته يتوقع أن تكون $b_3 > 0$ و $b_4 < 0$.

والآن، نحن في وضع يسمح لنا بمراجعة الأشياء. فإذا كان لدينا أربعة فصول واعتقدنا أن مستوى معادلتنا يتغير وفقا لكل فصل، نأخذ أحد هذه الفصول ونعده فصلنا المعياري. وننسب تأثير الفصول الأخرى لتأثيره. ففي المثال السابق اخترنا فصل الشتاء ليكون الفصل المعياري، أما إذا أسقطنا S_t وضمنا H_t ، F_t و W_t في معادلتنا، يصبح فصل الربيع، في هذه الحال، هو الفصل المعياري. وهنا، ستمثل b_0 ارتفاع المعادلة خلال فصل الربيع. وبمعنى آخر، يمكن القول إنه إذا كان لدينا أربعة فصول واعتقدنا بأن القاطع الرأسي لمعادلتنا يتغير وفقا لكل فصل، فإننا نكون معادلة تحتوي على أربع معلمات تصف هذه القواطع للمعادلة. ولما كان القاطع هو أحد هذه المعلمات فإننا نحتاج فقط إلى ثلاثة متغيرات صورية. لاحظ أننا لم نستطع تقدير المعلمات لمعادلتنا الأصلية (5.53) لأن هذه المعادلة تحتوي على خمس معلمات للقاطع هي: b_0 ، b_2 ، b_3 ، b_4 ، b_5 . ولكن، هناك فقط أربعة قواطع متعلقة بالفصول الأربعة. ولذا فإن أحد المتغيرات الصورية لاداعي له. ويصبح تعميما: إذا توقعنا أن التغيرات الإقليمية أو الفصلية تسبب k مستويات مختلفة من المعادلة فإننا نحتاج لاستخدام $(k-1)$ فقط من المتغيرات الصورية. وحتى تستوعب هذه المناقشة على نحو

أفضل، يمكنك العودة إلى معادلة الاستهلاك التي استخدمناها في بداية هذا المبحث، واعتبر ما يمكن أن يحدث فيما لو وضعنا في المعادلة (5.42) متغيراً صورياً لسنوات الحرب:

$$1 = W \quad \text{خلال سنوات الحرب}$$

$$0 \quad \text{خلال السلم،}$$

ومتغيراً صورياً ثانياً لسنوات السلم

$$1 = P \quad \text{خلال سنوات السلم،}$$

$$0 \quad \text{خلال أوقات الحرب.}$$

وبالتحديد، يجب أن تكون قادراً على إثبات أنه إذا أدخلنا كلا من W و P في المعادلة، فإنه لا يمكن تقديرها. كما ينبغي أن تبرهن بأن معادلة واحدة تغطي كلا من فترات الحرب والسلم يمكن تكوينها بأي من W أو P .

(٣-٥) الشكل الدالي مرة أخرى

في الفصل الثالث، فحصنا عدداً من التحويلات التي تمكنا من وضع العلاقات غير الخطية في شكل خطي، وذلك حتى يمكننا استخدام نموذج الانحدار الخطي العادي. يمكن تعميم استخدام هذه التحويلات بسهولة، ففي حالة الانحدار المتعدد، لن نتعمق في هذا الموضوع، وبدلاً من ذلك، سنأخذ مثلاً واحداً معنا - التحويل اللوغارتمي - لنرى كيف نوسع تحليلنا من حالة نموذج انحدار المتغيرين إلى حالة الانحدار المتعدد. وحينئذ، سنطور تحويلات إضافية أخرى مفيدة تصبح ممكنة عندما لانكون مقيدين بحالة انحدار المتغيرين السابق.

التحويل اللوغارتمي المعمم

تذكر من الفصل الثالث أننا اخترنا علاقة إنتاجية بسيطة من الشكل:

$$Q_t = aL_t^b e^{u_t}, \quad (5.57)$$

حيث L_t هي كمية العمل المستخدمة في الفترة t ، وهو عنصر الإنتاج الوحيد في إنتاج Q_t . وبأخذ اللوغاريتمات في المعادلة (5.57)، نعبّر عن دالة الإنتاج هذه على النحو:

$$\ln Q_t = \ln a + b \ln L_t + u_t. \quad (5.58)$$

بعدئذ، نحول المعادلة (5.58) إلى الشكل رقما خطي:

$$Q_t^* = a^* + bL_t^* + u_t, \quad (5.59)$$

عن طريق وضع:

$$Q_t^* = \ln Q_t,$$

$$a^* = \ln a,$$

$$L_t^* = \ln L_t.$$

في هذا الشكل، رقما استخدمنا منهجنا العادي للتقدير للحصول على المقدرات غير المتحيزة \hat{a}^* ، \hat{b} للمعلمات في المعادلة (5.59)، وبعدئذ، أخذنا $e^{\hat{a}^*}$ مقدراً متحيزاً ولكن متسقاً لـ a .

أحد القيود الواضحة في المعادلة (5.57) هو اشتغالها على عامل متغير واحد للإنتاج، ذلك أن السلع والخدمات تنتج، عادة، باستخدام تشكيلة من المدخلات. ويوحى هذا بأن دالة الإنتاج الأكثر واقعية ينبغي أن تشمل على كميات متغيرة من عوامل إنتاجية متعددة. لتحقيق ذلك دعنا نأخذ شكلاً عاماً للمعادلة (5.57):

$$Q_t = b_0 F_{1t}^{b_1} F_{2t}^{b_2} \dots F_{kt}^{b_k} e^{u_t}, \quad (5.60)$$

حيث يمثل كل من F_t 's كمية من عامل إنتاجي معين يستخدم خلال الفترة t . على سبيل المثال، قد تشير F_{1t} إلى كمية العمل المستخدمة في الفترة t ، بينما تشير F_{2t} إلى

* تحتوي دالة الإنتاج لكوب ودوجلاس على عدد من السمات الملائمة والمهمة جعلتها ذات استخدام كبير في التحليل الاقتصادي. لمناقشة هذه السمات، يرجع إلى:

James M. Henderson and Richard E. Quandt, *Micro-economic Theory*, 2nd ed. (New York: McGraw-Hill, 1971), pp. 79-85.

مساحة الأرض وهلم جرا. وعلاقة الإنتاج التي تأخذ مثل هذا الشكل رقمتمعرف بدالة إنتاج كوب - دوجلاس Cobb-Douglas*. ومانحتاج معرفته هو قيم b 's في هذه العلاقة حتى يمكننا أن نعرف كيف يتغير الناتج الكلي مع تغير كميات المدخلات المختلفة. فقد نكون مهتمين، على سبيل المثال، بمعرفة ما إذا كان إنتاج سلعة معينة يخضع لظاهرة تزايد غلة الحجم. ومعنى ذلك أنه مع ثبات الأشياء الأخرى^{*}، إذا استطعنا مضاعفة المدخلات من جميع عناصر الإنتاج الأخرى، هل يزداد الإنتاج بأكثر من الضعف؟ فإذا كان الأمر كذلك فإننا نصرح بأن لدينا تزيادا في غلة الحجم، أما إذا تزايد الإنتاج إلى الضعف بالضبط، حيثذ هناك ثبات في غلة الحجم، وأخيرا إذا تزايد الإنتاج بأقل من الضعف، يوجد لدينا حيثذ تناقصا في غلة الحجم.

ويمكن تحديد ذلك بسهولة من دالة إنتاج كوب-دوجلاس عن طريق أخذ المجموع $\sum_{i=1}^n b_i$. ويتضح لنا ذلك بأخذ مثال مبسط، ونترك التعميم للقارئ على سبيل التمرين. افترض أن لدينا سلعة، Q ، تنتج باستخدام عنصري العمل ورأس المال فقط على النحو:

$$Q_t = b_0 L_t^{b_1} K_t^{b_2} e^{u_t}, \quad (5.61)$$

حيث L_t و K_t هما كميات العمل ورأس المال، على التوالي والمستخدم في إنتاج Q خلال الفترة t . والآن افترض أننا ضاعفنا مدخلات العمل ورأس المال، دع Q'_t هو المستوى الجديد من الإنتاج، حيثذ، سيكون لدينا:

$$\begin{aligned} Q'_t &= b_0 (2L_t)^{b_1} (2K_t)^{b_2} e^{u_t} = b_0 (2^{b_1}) (L_t^{b_1}) (2^{b_2}) (K_t^{b_2}) e^{u_t} \\ &= 2^{(b_1+b_2)} b_0 L_t^{b_1} K_t^{b_2} e^{u_t} = 2^{(b_1+b_2)} Q_t. \end{aligned} \quad (5.62)$$

يتضح لنا من المعادلة الأخيرة في (5.62) أنه إذا كانت لدينا $(b_1 + b_2) > 1$ فإن الإنتاج سوف يزداد بأكثر من الضعف وسيكون لدينا تزايد في غلة الحجم، وإذا كانت $b_1 + b_2 = 1$ يتضاعف الإنتاج بالضبط، فهناك ثبات في غلة الحجم، وأخيرا

* يرتبط الشرط «مع بقاء الأشياء الأخرى على حالها» بالخطأ العشوائي، أي أننا نفترض، في المناقشة أن الخطأ العشوائي لا يتغير عندما تتغير قيم المدخلات.

إذا كانت $b_1 + b_2 < 1$ فإن الإنتاج يزداد بأقل من الضعف حيث يوجد تناقص في غلة الحجم. وبعمومية أكثر في المعادلة (5.60) ينبغي أن تكون قادراً على إثبات أن حالات تزايد غلة الحجم وثباتها وتناقصها تناظر الحالات التي يزيد فيها $\sum_{i=1}^n b_i$ عن الواحد، يساوي الواحد أو يقل عن الواحد الصحيح على التوالي*. وعملياً تصبح مشكلتنا هي تقدير قيمة الـ b 's من أجل تحديد طبيعة العلاقة الإنتاجية لسلعة معينة. وللحصول على المعادلة (5.60)، في شكل قابل للتقدير نستخدم التحويل اللوغارتمي، فبأخذ لوغاريتمات المعادلة (5.60)، نحصل على:

$$\ln Q_t = \ln b_0 + b_1 \ln F_{1t} + b_2 \ln F_{2t} + \dots + b_k \ln F_{kt} + u_t. \quad (5.63)$$

بعدئذ، نعرف:

$$Q_t^* = \ln Q_t,$$

$$b_0^* = \ln b_0,$$

$$F_{it}^* = \ln F_{it}.$$

وبالتعويض في المعادلة (5.63)، يصبح لدينا مايلي:

$$Q_t^* = b_0^* + b_1 F_{1t}^* + b_2 F_{2t}^* + \dots + b_k F_{kt}^* + u_t. \quad (5.64)$$

وباستخدام طريقتنا العادية في التقدير على المتغيرات المعرفة في المعادلة (5.64)، يمكننا الحصول على مقدرات غير متحيزة \hat{b}_0 ، \hat{b}_1 ، ... و \hat{b}_k لمعاملات المعادلة (5.64)، وبهذه الطريقة نحصل على مقدرات غير متحيزة لمكونات الناتج لكل عنصر من عناصر الإنتاج. وكما هو في حالة الانحدار البسيط، سيكون $\hat{b}_0 = e^{\hat{b}_0}$. وأخيراً، وباستخدام النتائج الموضحة في الملحق B لهذا الفصل، يمكننا اختبار فرضية وجود ثبات غلة الحجم.

* إحدى السمات الأخرى المفيدة لدالة الإنتاج كوب - دوجلاس في المعادلة (5.60) هي أن كل b_i يمكن تفسيرها على أنها مرونة الناتج بالنسبة للعامل i . أي إذا كانت F_i تزايد بنسبة ١٪، وجميع المدخلات الأخرى تظل ثابتة، فإن الناتج Q سيزداد بنسبة b_i في المائة. لكن، على القارئ أن يلاحظ أن كل متغير لا يمكن الاستغناء عنه في عملية الإنتاج، بمعنى أنه إذا كانت $F_i = 0$ ، فإن الناتج Q سيصبح صفراً أيضاً.

أشكال متعددة الحدود للمتغيرات المستقلة

يود الاقتصاديون، عادة، التعامل مع إمكانية أن تكون العلاقات بين المتغيرات الاقتصادية غير خطية إلا أنهم غير متأكدين، في الحقيقة، من شكل هذه العلاقة. فمثلاً، اعتبر تأثير عمر الفرد على إنفاقه الاستهلاكي. فمن الممكن أنه، مع تقدم عمر الفرد واتساع خبراته، أن تحفره معلوماته عن الأنشطة المختلفة على زيادة إنفاقه على السلع والخدمات الاستهلاكية. ولكن، بعد وصوله إلى عمر معين، قد يتباطأ إنفاق الفرد، وفي الحقيقة، يبدأ الفرد في تخفيض مستوى إنفاقه الاستهلاكي. مثل هذه العلاقة بين الإنفاق الاستهلاكي والعمر (عندما تظل المتغيرات الأخرى الملائمة كمستوى الدخل ثابتة) تظهر في المنحنى المتصل A في الشكل رقم (٥-٩).

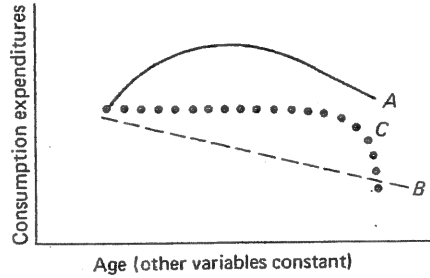
من ناحية، أخرى قد نجد أنه، مع تقدم الفرد في العمر، يزداد طلبه على الأمن الاقتصادي ومن ثم، على الادخار، ولذا، يتناقص إنفاقه الاستهلاكي بانتظام. مثل هذه العلاقة تظهر في الخط المتقطع B في الشكل نفسه. وأخيراً قد يتناقص إنفاق الفرد الاستهلاكي في البداية ببطء شديد وبعدئذ، كلما تقدم الفرد في العمر يتناقص استهلاكه بمعدل متزايد. وتظهر هذه العلاقة في الخط المنقط C في الشكل رقم (٥-٩).

وعلى سبيل أحد الأمثلة الأخرى للمتغيرات التي ترتبط بعضها بعضاً بعلاقة أقرب إلى عدم الخطية، اعتبر تأثير التغيرات في تكلفة المعيشة على تعديلات الأجور، كما تقاس بالتغيرات في الأجور النقدية. وبالطبع يمكن جعل التغيرات في تكلفة المعيشة تنعكس تماماً في تغيرات الأجور. أي أنه، مع بقاء الأشياء الأخرى، على حالها، إذا ارتفعت تكلفة المعيشة بمقدار X في المائة ارتفع معدل الأجور بمقدار X في المائة*. ومن الناحية الأخرى من المحتمل أن التغيرات الطفيفة في تكلفة المعيشة لاتلاحظ ومن ثم لاتؤدي إلى زيادات مناظرة في الأجور. في هذه الحال، قد نفترض أن التغيرات الكبيرة في تكلفة المعيشة هي فقط التي تنعكس في تعديلات الأجور. تظهر هذه الاحتمالات في منحنيات OA_1 و OA_2A_3 على الترتيب في الشكل رقم (٥-١٠).

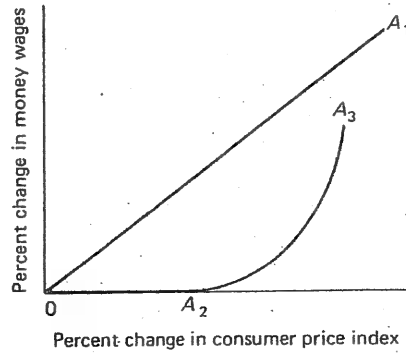
* قد ترغب في إضافة بعض الزيادات الإضافية في الأجور لتعكس الزيادة في الإنتاجية.

وفي ضوء هذه الأمثلة، نرغب في توجيه اهتمامنا نحو مشاكل التقدير واختبار العلاقة بين المتغيرات عندما لا يكون شكل هذه العلاقة مؤكداً. سنقيم نتائجنا من خلال عرض أمثلة إضافية في مبحث آخر من هذا الفصل. نبدأ (استخدام حالة المتغيرين للتبسيط) بافتراض أن المتغير التابع Y يرتبط بالمتغير المستقل X ارتباطاً غير مؤكد. هذا الافتراض يمكن التعبير عنه على النحو:

$$Y_i = f(X_i) + u_i, \quad (5.65)$$



شكل (٩-٥)



شكل (١٠-٥)

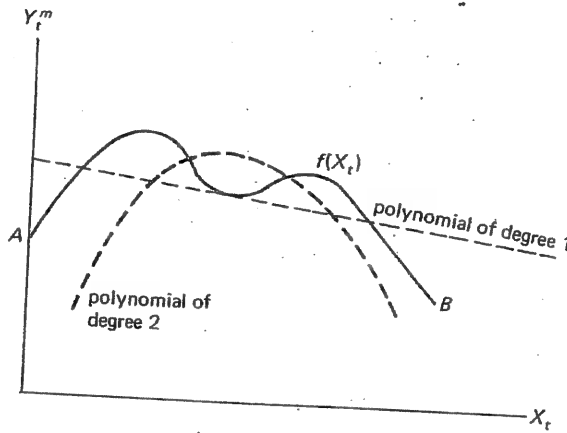
حيث إن u_t هو الخطأ العشوائي. تبين المعادلة (5.65)، ببساطة أن القيمة رقم t لـ Y أو Y_t ، تعتمد على القيمة رقم t لـ X أو X_t ، وعلى الخطأ العشوائي u_t . ولأننا لانعلم الشكل المحدد لـ $f(x_t)$ في المعادلة (5.65)، فإنه لتحقيق بعض التقدم في تقدير العلاقة بين Y_t و X_t ينبغي علينا إما أن نوجد شكل $f(x_t)$ أو بالمقابل، استخدام بعض التقريب لها. سوف نستخدم المنهج الأخير عن طريق استدعاء النظرية التي ذكرناها واستخدمناها عند الحديث عن طريقة فترة الإبطاء لآلمون. وبالتحديد، تصرح النظرية، أنه في ظل ظروف عامة، فإن دالة (أو منحنى) قد تقرب لأي درجة من الدقة بواسطة متعدد للحدود. فإذا كان الأمر كذلك فيمكننا أن نطبق هذه النظرية على دالتنا غير المعلومة في المعادلة (5.65) للحصول على:

$$f(X_t) = a_0 + a_1 X_t + a_2 X_t^2 + \dots + a_k X_t^k. \quad (5.66)$$

وعموماً، كلما أردنا درجة أكبر من الدقة ينبغي علينا استخدام درجة أعلى من متعدد الحدود (k). ينتج هذا من مناقشتنا بمتعدد الحدود في طريقة فترة الإبطاء لآلمون. في ذلك المبحث، اشرنا إلى أن عدد نقاط الانقلاب على شكل متعدد الحدود تقل عن درجة متعدد الحدود بمقدار واحد على الأكثر. من هذا، قد نستخلص أن متعدد الحدود ذا الدرجات الأعلى يكون أكثر مرونة من ذلك المرتبط بدرجات أقل. ويشير هذا إلى أننا إذا أردنا تقريبا أدق فإننا نحتاج إلى متعدد للحدود ذي درجة أعلى حتى يتضمن مرونة كافية تتبع اتباعاً وثيقاً شكل الدالة غير المعلومة. وبالتدريج، فكلما كان شكل الدالة التي نرغب في تقريبها معقداً ازدادت بالضرورة درجة متعدد الحدود المناظر.

وقد وضح ذلك بدلالة الدالة غير المعلومة $f(x_t)$ في الشكل رقم (٥-١١). في هذا الشكل نفترض أن دالتنا غير المعلومة لها الشكل المشار إليه بالخط المتصل AB. الآن يمكن تعريف هذا المنحنى - على الرغم من أن ذلك يتم بطريقة ضعيفة - بخط مستقيم يترتب على استخدام متعدد للحدود من الدرجة الأولى ($k=1$). ويتحسن التقريب إذا ما استخدمنا بدلاً من ذلك متعدد للحدود من الدرجة الثانية ($k=2$). ويمكن أن يتحسن التعريف أكثر عن طريق $k=3$ وهلم جرا.

افترض أنه يتوافر لدينا عدد من الافتراضات المرتبطة بالشكل رقم العام للعلاقة بين Y_t و X_t . افترض، أيضا، أن أكثر هذه الافتراضات تعقيدا تقترح أن قيمة $k = k_m$. ستكون ملائمة للتقريب. في المعادلة (5.66)*، ونعني بكلمة «ملائمة» أن علامة تقريبا مساوية لـ = في المعادلة (5.66) قد تستبدل بخسارة قليلة من الدقة بعلامة التساوي. نلاحظ، أيضا، أنه إذا كان متعدد الحدود من الدرجة k_m هو تقريب ملائم لأكثر أشكالنا المفترضة تعقيدا، فإنه يكون تقريبا ملائما، أيضا، للأشكال الأبسط كافة التي نعتبرها.



شكل (١١-٥)

في ظل هذا الافتراض المرتبط بـ k_m ، قد نستخدم $k = k_m$ في المعادلة (5.66) وهذا، بدوره، في المعادلة (5.65) للحصول على:

$$Y_t = a_0 + a_1 X_t + a_2 X_t^2 + \dots + a_{k_m} X_t^{k_m} + u_t. \quad (5.67)$$

ويمكن تحويل المعادلة (5.67) إلى نموذجنا العادي عن طريق التعويضات:

$$Z_{it} = X_t^i, \quad i = 1, \dots, k_m \quad (5.68)$$

أي إذا قمنا بالتعويض من المعادلة (5.68) في المعادلة (5.67) سنحصل على:

$$Y_t = a_0 + a_1 Z_{1t} + a_2 Z_{2t} + \dots + a_{k_m} Z_{k_m t} + u_t, \quad (5.69)$$

* بالنسبة لمعظم التطبيقات الاقتصادية، تعد قيمة $k = 3$ معقولة.

والذي يمثل الشكل العادي.

دع $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_{k_m}$ هي مقدرات المعلمات للمعادلة (5.69)، حيث أن العلاقة المقدرة بين Y_t و X_t هي:

$$\hat{Y}_t = \hat{a}_0 + \hat{a}_1 X_{t1} + \dots + \hat{a}_{k_m} X_{t k_m}. \quad (5.70)$$

لأن مقدرنا $f(x_t)$ سيحصل عليه بوساطة:

$$\hat{f}(X_t) = \hat{a}_0 + \hat{a}_1 X_{t1} + \dots + \hat{a}_{k_m} X_{t k_m}. \quad (5.71)$$

اعتبر الآن مشكلة ما إذا كان المتغير Y_t يعتمد على X_t أم لا. عند النظر الأولى قد يظهر أن هذا الافتراض يمكن اختباره ببساطة عن طريق اختبار الفرضيات، واحدا بعد الآخر، بأن $a_1=0, a_2=0, \dots, a_{k_m}=0$ في المعادلة (5.64). وسوف نخلص افتراضا إلى أن Y_t و X_t مرتبطان بقوة ببعضهما بعضا عند رفض أي من فرضيات العدم هذه. وعلى العكس، إذا قبلنا جميع فرضيات العدم هذه ستصبح النتيجة هي أن Y_t و X_t غير مرتبطتين بقوة ببعضهما بعضا.

ولسوء الحظ، لا يمكن اختبار هذه الفرضيات المرتبطة بعلاقة بين Y_t و X_t بهذه الطريقة بسبب ما يمكن أن يسمى «خدعة التجميع» «Fallcy of Composition». أي أن الفرضية في هذه الحال، ترتبط بأكثر من معلمة. بخاصة، أن الفرضية التي نرغب في اختبارها هي:

$$H_0: a_1 = a_2 = \dots = a_{k_m} = 0. \quad (5.72)$$

ولقد طورنا في ملحق هذا الفصل طريقة لاختبار الفرضيات على شاكلة المعادلة (5.75)، ولكننا نشير عند هذه النقطة إلى أنه، إذا كان $a_1=0, a_2=0, \dots, a_{k_m}=0$ قد اختيرت واحدة تلو الأخرى عند مستوى معين من المعنوية، مثلا 5% وقبلت فإن امكانية رفض الفرضية (5.72) لاتزال عند المستوى نفسه من المعنوية. وبمعنى آخر، فإن الفرضية التي ترتبط بأكثر من معلمة، مثلا k_m ، كما في المعادلة (5.72) لا يمكن، عموما، اختبارها على نحو عام عند مستوى معنوية عن طريق تقسيمها إلى فرضيات عددها k_m كل منها يرتبط بمعلمة مفردة ثم اختبار هذه الفرضيات

على التوالي عند ذلك المستوى من المعنوية.

إن توسيع طريقة التقدير أعلاه للحالة التي يشتمل فيها نموذج الانحدار على متغيرات مستقلة إضافية سهل ومباشر. افترض مثلاً أن نموذجاً من الشكل:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + f(X_{3t}) + u_t, \quad (5.73)$$

حيث إن u_t ، مرة أخرى، هي الخطأ العشوائي، وأتينا نفترض، أيضاً، أن الشكل رقمالمعين لـ $f(X_{3t})$ غير مؤكد. حينئذ، يترتب على منهجنا أعلاه أنه إذا كانت فرضياتنا المرتبطة بـ $f(X_{3t})$ توحى بأن $k = k_m$ سيعطي تقريباً ملائماً لمتعدد الحدود، فسوف نفترض أن:

$$f(X_{3t}) = a_0 + a_1 X_{3t} + \dots + a_{k_m} X_{3t}^{k_m}. \quad (5.74)$$

وبالتعويض من المعادلة (5.74) في المعادلة (5.73) يتتج:

$$Y_t = A + b_1 X_{1t} + b_2 X_{2t} + a_1 X_{3t} + \dots + a_{k_m} X_{3t}^{k_m} + u_t, \quad (5.75)$$

حيث إن $A = a_0 + b_0$ ، مرة أخرى بجعل:

$$Z_{it} = X_{3t}^i, \quad i = 1, 2, \dots, k_m, \quad (5.76)$$

فإنه يمكن كتابة المعادلة (5.75) على النحو:

$$Y_t = A + b_1 X_{1t} + b_2 X_{2t} + a_1 Z_{1t} + \dots + a_{k_m} Z_{k_m t} + u_t, \quad (5.77)$$

والتي هي في الشكل العادي. ويترتب على ذلك أنه - وبافتراض إمكانية تقريب متعدد الحدود - يمكننا الحصول على مقدرات غير متحيزة لـ \hat{A} ، \hat{b}_1 ، \hat{b}_2 ، \hat{a}_1 ، \dots ، \hat{a}_{k_m} للمعاملات في المعادلة (5.77).

لاحظ أننا، في هذه الحال، لن نقدر على الحصول على مقدرات منفصلة لـ a_0 و b_0 لأن مجموع A هو الذي يظهر في المعادلة (5.77). وبخلاف الحال، المبسطة أعلاه، سنقدر على تقدير $f(X_{3t})$ حتى ثابت تجميعي additive constant، وبمعنى آخر سنقدر على تقدير الجانب المتغير لـ $f(X_{3t})$ ، مثلاً $f_v(X_{3t})$:

$$\widehat{f_v(X_{3t})} = \hat{a}_1 X_{3t} + \dots + \hat{a}_{k_m} X_{3t}^{k_m}. \quad (5.78)$$

وعادة ما يكون الجزء المهم من $f(X_{3t})$ هو جزؤه المتغير، بسبب أن هذا الجزء يصف الطريقة التي يتغير بها Y_t مع X_{3t} .
في هذه المرحلة ينبغي أن يتضح أن بإمكاننا توسيع الطريقة أعلاه لتشتمل عموماً، على حالة وجود أي عدد من المتغيرات المستقلة. وينبغي أن يكون واضحاً أيضاً أنه يمكن توسيعها لتشتمل على نموذج بأكثر من متغير مستقل يدخل بطريقة غير محددة على النموذج.*

توليفات من الأشكال الدالية

لتجميع مناقشتنا حول الأشكال الدالية، نؤكد أنه من المشروع تماماً استخدام أشكال عديدة من التحويلات المختلفة في معادلة الانحدار نفسها. وفي الحقيقة، قد نلاحظ في بعض امثلتنا السابقة، أن واحداً أو أكثر من المتغيرات تظهر في شكل لوغاريتمي أو ربما في شكل عكسي، بينما لا تكون المتغيرات الأخرى قابلة لأي شكل من التحويلات بتاتا. وعلى سبيل مثال إضافي، اعتبر الشكل التالي الأكثر تعقيداً لعلاقة منحنى فليبيس:

$$\dot{W}_t = b_0 + b_1 \left(\frac{1}{R_t} \right) + b_2 \pi_{(t-1)} + b_3 \dot{P}_t + b_4 \dot{P}_t^2 + b_5 \ln G_t + u_t, \quad (5.79)$$

حيث:

\dot{W} = التغير النسبي للأجور خلال الفترة t ,

R_t = معدل البطالة في الفترة t ,

$\pi_{(t-1)}$ = معدل الأرباح للمنشآت في الفترة $t-1$,

\dot{P}_t = التغير النسبي للأسعار في الفترة t ,

$\ln G_t$ = اللوغارتم الطبيعي لمعدل النمو في قوة العمل في الفترة t ,

u_t = الخطأ العشوائي.

* اعتبر على سبيل المثال، نمودجا يأخذ الشكل:

$$Y_t = a_0 + f_1(X_{1t}) + f_2(X_{2t}) + a_1 X_{3t} + u_t.$$

لاحظ أننا سنستخدم في المعادلة نفسها التحويل العكسي، التحويل اللوغاريتمي، علاقة مبطأة، وشكلا لمتعدد الحدود لواحد من المتغيرات المستقلة. والآن سنعتبر المعادلة (5.79) لكي نوضح طرق التحويل، ولكن، عند التطبيق، تكون هناك عادة «أسباباً» (فرضيات اقتصادية) وراء كل تحويل. فمثلاً رأينا في الفصل الثالث أن علاقة خطية بين W و R ليست ملائمة تماماً (بسبب أن R لا يمكن أن يأخذ قيما سالبة)، وأن التحويل العكسي ربما يكون منطقياً. وبالنسبة لمعدل الأرباح $\pi_{(t-1)}$ قد نتوقع أن اتحادات العمال التي تقوم بالمفاوضات (حول معدلات الزيادة في الأجور) تستخدم معدل الأرباح (والذي يكون عادة متاحاً للفترة المحاسبية السابقة) باعتباره أحد العناصر في عملية المفاوضة. فإذا كانت الأرباح في الفترة الأخيرة عالية بصورة غير عادية فإن اتحاد العمال قد يجد مبرراً في المطالبة بزيادات مناظرة أكبر في الأجور. وبالمثل قد يتغير شكل متعدد الحدود لتغير السعر، لأن العلاقة بين تعديلات الأجور وتعديلات الأسعار قد تكون غير خطية وذلك إذا كانت التغيرات السعريّة الطفيفة تحدث دون أن يلاحظها العمال بينما تكون التغيرات السعريّة الكبيرة مؤشراً للمطالبات بالزيادات في الأجور (انظر الشكل رقم ٥-١٠). وأخيراً، ولأن النمو في قوة العمل يرتبط بزيادة عرض العمل، فقد يكون له تأثير على المركز التفاوضي في قوة العمل. وفي هذا المجال، رأينا في الفصل الثالث أن المتغير الأكثر ملائمة هو الزيادة النسبية (وليس المطلق)، لذلك فإن حجم التغير يقاس بالنسبة إلى المستوى الحالي لعرض العمل. وللتوضيح سنستخدم في المعادلة (5.79) التحويل اللوغاريتمي للمتغير G_t . والنقطة المهمة وراء كل هذا هي ببساطة أن اختيار الشكل الدالي ليس أساساً عملية تجريبية وخطأ، وإنما ينبغي استخدام المعلومات النظرية الاجتماعية المتاحة لدينا في تحديد الشكل الأكثر ملاءمة لعلاقتنا الدالية.

وبالعودة إلى معادلة الأجور، يمكننا وضع المعادلة (5.79) في شكل خطي

باستخدام التحويلات التالية:*

$$Z_{1t} = \left(\frac{1}{R_t} \right),$$

$$Z_{2t} = \pi_{(t-1)},$$

$$Z_{3t} = \dot{P}_t,$$

$$Z_{4t} = \dot{P}_t^2,$$

$$Z_{5t} = \ln G_t.$$

وبالتعويض عن هذه التحويلات في المعادلة (5.79)، نحصل على الشكل الخطي لمعادلتنا على النحو التالي:

$$\dot{W}_t = b_0 + b_1 Z_{1t} + b_2 Z_{2t} + b_3 Z_{3t} + b_4 Z_{4t} + b_5 Z_{5t} + u_t. \quad (5.80)$$

ويمكننا ببساطة حساب قيمة Z's من القيم المشاهدة لكل من R ، π ، \dot{P} و G ومن ثم باستخدام هذه القيم Z's وطريقتنا العادية لتقدير يمكن حساب \hat{b}_0 ، \hat{b}_1 ، \hat{b}_2 ، \hat{b}_3 ، \hat{b}_4 و \hat{b}_5 .

كل هذا ينبغي أن يوضح مدى المرونة الموجودة في تقدير الانحدار المتعدد. كثيرا ما أنتقد الاقتصاديون القياسيون خاصة في بداية استخدام تحليل الانحدار بسبب ازدياد درجة اعتمادهم على الشكل الخطي من العلاقات، ولكن ينبغي الآن أن تكون قادرا على رؤية أن الاستخدام المنطقي، والذكي، للتحويلات يجعل نموذج الانحدار المتعدد قادرا على تناول تشكيلة عريضة من الأشكال الدالية المعقدة.

(٤-٥) توضيح: الطلب على النقود

إحدى القضايا المركزية في اقتصاديات النقود هي نظرية الطلب على النقود وقياسها*. في الحقيقة فإن كثيرا مما يتعلق بهذه القضية، كالتأثير المحتمل للسياستين

* في الحقيقة، فإن التحويل $Z_{3t} = \dot{P}_t$ غير ضروري وقد وضع، فقط، لتحقيق الاتساق في وضع الرموز.

المالية والنقدية على النشاط الاقتصادي يعتمد على شكل دالة الطلب على النقود وعلى قيم معلوماتها. وتفيد النظرية الاقتصادية أن الطلب على الأرصدة النقدية الحقيقية (أي الأرصدة النقدية الرسمية معدلة بالمستوى العام للأسعار وذلك لتثبيت قوتها الشرائية) يعتمد على ثلاثة أنواع على الأقل من المتغيرات: الدخل، معدل الفائدة على السندات (أو معدل العائد على الأصول المالية الأخرى) وربما صافي الثروة. باختصار فإنه مع زيادة دخل الأفراد يزداد طلبهم على الأرصدة النقدية لأغراض التبادل، ومع ارتفاع أسعار الفائدة ينخفض طلبهم على الأرصدة النقدية لارتفاع تكلفة الفرصة البديلة الفعالة للاحتفاظ بالأرصدة النقدية (والتي، لم تكن تحقق فائدة في الأقل، إلا مؤخرا)، ومع ارتفاع ثروات الأفراد فإنهم سوف يميلون للاحتفاظ بقدر أكبر من الأرصدة النقدية كأحد أشكال الاحتفاظ بالثروة المتزايدة. يمكننا تلخيص كل هذا في المعادلة:

$$M_d = f(Y, r, W), \quad (5.81)$$

حيث:

$$M_d = \text{الطلب على الأرصدة النقدية،}$$

$$Y = \text{الدخل الحقيقي،}$$

$$r = \text{معدل الفائدة، وأخيرا،}$$

$$W = \text{الثروة (أو صافي الثروة).}$$

حيث نتوقع أن يكون التأثير الجزئي لسعر الفائدة سالبا أما التأثير الجزئي لمتغيرات الدخل والثروة فتتوقع أن يكون موجبا.

قام عدد من الاقتصاديين ببحوث قياسية مكثفة بهدف تقدير معادلات تناظر المعادلة (5.81)، اعتمدت أعمالهم على مجموعة من الأشكال الدالية المتضمنة في

* لمعالجة أكثر تفصيلا يرجع إلى:

David E. Laidler, *The Demand for Money: Theories and Evidence*, 2nd ed., (New York: Dun-Donnelley, 1977).

عديد من التحويلات التي اعتبرناها في هذا الفصل. وتشتمل هذه، أيضا، على قيم مبطأة لبعض المتغيرات. وعلى سبيل المثال، نعرض هنا نتائج إحدى الدراسات (قام بها Martin Bronfenbrenner and Thomas Mayer)*. ونقطة البداية لديهما هي وضع دالة الطلب على النقود على شكل حاصل ضرب كالتالي:

$$M_{dt} = b_0 Y_t^{b_1} r_t^{b_2} W_t^{b_3} M_{d(t-1)}^{b_4} e^{u_t}, \quad (5.82)$$

حيث u_t هو الخطأ العشوائي، وأن جميع المتغيرات الأخرى قد عرفت من قبل في المعادلة (5.81)، وبأخذ اللوغاريتمات للجانبين المعادلة (5.82) استطاع الباحثان أن يحصلوا على الشكل الخطي التالي:

$$\ln M_{dt} = \ln b_0 + b_1 \ln Y_t + b_2 \ln r_t + b_3 \ln W_t + b_4 \ln M_{d(t-1)} + u_t. \quad (5.83)$$

ولأغراض التفسير، افترض أن u_t تحقق افتراضاتنا العادية كافة، ونعيد كتابة المعادلة (5.83) على النحو التالي:

$$(\ln M_{dt} - b_4 \ln M_{d(t-1)}) = B + b_1 \ln Y_t + b_2 \ln r_t + b_3 \ln W_t + u_t. \quad (5.84)$$

حيث إن $B = \ln b_0$. في هذا الشكل يمكن أن يستخدم نموذج مثل المعادلة (5.83) مفسراً للفرق بين القيمة الحالية للمتغير التابع ($\ln M_{dt}$) وقيمته المبطة والمضروبة في المعامل b_4 . فمثلا افترض أننا نشعر بأن لوغاريتم الطلب على النقود يتقلب عشوائيا حول خط اتجاه يتزايد بمعدل ٣٪ مثلا. في هذه الحال، فإن توقعاتنا المتعلقة بقيم لمعاملات في المعادلة (5.83) أو المعادلة (5.84) هي $b_4 = 1.03$ و $B = b_1 = b_2 = b_3 = 0$. وسنعتقد بأن الطلب على النقود لا يستجيب لمتغيرات الدخل وسعر الفائدة والثروة.

باستخدام بيانات سنوية للولايات المتحدة الأمريكية للفترة ١٩١٩ - ١٩٥٦م، قدر الباحثان المعادلة (3.83) بطريقة المربعات الصغرى وحصلوا على:

* يرجع إلى:

Martin Bronfenbrenner and Thomas Mayer. "Liquidity Functions in the American Economy." *Econometrica*, 28 (1960), pp. 810-834.

$$\ln M = 0.11 + 0.34 \ln Y - 0.09 \ln r - 0.12 \ln W + 0.72 \ln M_{t-1} \quad (5.85)$$

(0.03) (0.09) (0.01) (0.08) (0.06)

$R^2 = 0.99,$

حيث تمثل الأرقام داخل الأقواس أسفل معاملات الانحرافات المعيارية المقدرة المناظرة. ولتوضيح النتائج، فإن قيم t المحسوبة والمناظرة لمتغيرات الدخل، معدل الفائدة، الثروة، والمتغير المبطل للطلب على النقود هي على الترتيب 9.8، 1.5، و12 تقريباً. وعندما نستخدم قاعدتنا التجريبية، تدل هذه النتائج على أنه إذا اعتبرنا أن فرضية العدم $H_0: b_3 = 0$ مقابل الفرضية البديلة $H_1: b_3 \neq 0$ عند مستوى معنوية 5% فإننا سنقبل فرضية العدم. وبالمقابل إذا افترضنا أي فرضية من فرضيات العدم الأخرى $H_0: b_i = 0$ مقابل $H_1: b_i \neq 0$ ، حيث $i = 1, 2, 4$ عند مستوى معنوية 5% فسنرفض فرضية العدم.

لاحظ أن فرضية العدم $H_0: b_2 = 0$ لها أهمية خاصة عند الاقتصاديين النقديين. ذلك أن رفض هذه الفرضية سوف يمنحهم سبباً للاعتقاد بأنه عند معدلات الفائدة الأعلى يحتفظ الأفراد بنسبة أصغر من ثرواتهم في شكل أرصدة نقدية حتى يستطيعون الاستفادة من العائد الأعلى المتاح من السندات ومن الأصول التمويلية الأخرى التي تعطي فائدة. أحد المتضمنات المهمة لهذه النتيجة هي أن السياسة المالية لها بعض التأثير على الطلب الكلي، فإذا لم يكن الطلب على النقود مستجيباً لسعر الفائدة فإن السياسة المالية سوف تحفز فقط تغيرات معوضة offsetting في الإنفاق الخاص بدون أي تأثير صافي على الإنفاق في الاقتصاد.*

* في هذه النقطة، انظر Laidler، المرجع نفسه، الفصل الثاني. باختصار، إذا كان الطلب على النقود لا يتأثر بالتغيرات في معدل الفائدة، فإن المنحنى LM سيكون رأسياً. ونتيجة لذلك، فإن التخفيض الضريبي أو الزيادة في معدل الإنفاق الحكومي سيدفع معدلات الفائدة إلى أعلى ومن ثم يقلل الإنفاق الخاص بكمية مساوية. في هذه الحال، فإن مستوى الناتج القومي الإجمالي يعتمد تماماً على عرض النقود.

ملحق أ (A): قيود طرفية في إبطاء ألون

يعالج هذا الملحق استخدام قيود طرفية مع منهج ألون لتقدير الانحدار ذي فترات الإبطاء، وكما ذكرنا في هذا الفصل من قبل، فقد يشعر الاقتصادي بأنه يعلم ليس، فقط، نمط b 's ولكنه يعلم، بالضبط، أيضا، قيمة أي من b_0 و b_k أو كليهما، وهذه القيمة عادة صفر. فإذا علمنا هذه القيم، فإنه ينبغي علينا محاولة إدخال هذه المعلومات في عملية تقديرنا لنموذج الانحدار.

افترض، على سبيل المثال، أننا نشعر بأن نمط b 's يكون مشابها لذلك المنحنى (A) الموجود في الشكل رقم (٥-أ) حيث تتناقص في البداية قيم b 's ثم تزايد، وأخيرا تتناقص حتى تصبح b_k مساوية للصفر. يمكننا في هذه الحال، أن نضع قيودا طرفيا واحدا وهو $b_k = 0$. فإذا افترضنا من الناحية الأخرى أن نمط b 's يأخذ شكل المنحنى (B) في الشكل رقم (٥-ب) فإنه في هذه الحال، يمكن أن يوجد قيدان طرفيان وهما $b_0 = b_k = 0$. وعموماً، يمكن لواضع النموذج فرض أي من القيود الطرفية السابقة أو كلها، وهذا المنهج هو تعميم مباشر لطرق التقدير المعروفة، افترض، مرة أخرى، وجود المعادلة (5.24)، وأن $k=10$ (أي أن معادلتنا تحتوي على عشر فترات إبطاء):

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + \dots + b_{10} X_{t-10} + u_t. \quad (5A.1)$$

افترض أولا (متجاهلا القيود الطرفية) أن النمط المقترح b 's يأخذ شكل متعدد الحدود من الدرجة الثالثة:

$$b_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \alpha_3 i^3, \quad i = 0, \dots, 10. \quad (5A.2)$$

حيث، بالتعويض من المعادلة (5A.2) لكل واحدة من b في المعادلة (5A.1)، نتج لنا المعادلة:

$$Y_t = a + \alpha_0 Z_{1t} + \alpha_1 Z_{2t} + \alpha_2 Z_{3t} + \alpha_3 Z_{4t} + u_t. \quad (5A.3)$$

حيث:

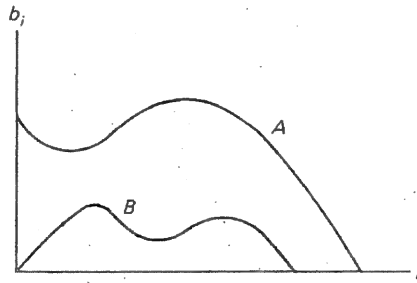
$$\begin{aligned} Z_{1t} &= \sum_{i=0}^{10} X_{t-i}, & Z_{2t} &= \sum_{i=1}^{10} i X_{t-i}, \\ Z_{3t} &= \sum_{i=0}^{10} i^2 X_{t-i}, & Z_{4t} &= \sum_{i=1}^{10} i^3 X_{t-i}, \end{aligned}$$

دعنا الآن نفرض قيدنا الطرفي، وبالتحديد افترض أننا نعتقد أن $b_{10} = 0$ حيثئذ فممن المعادلة (5A.2)، يكون لدينا:

$$b_{10} = \alpha_0 + 10\alpha_1 + 100\alpha_2 + 1000\alpha_3 = 0. \quad (5A.4)$$

ومن المعادلة (5A-4) يتضح أنه إذا فرضنا الشرط $b_{10} = 0$ فإنه ينبغي أن يكون لدينا:

$$\alpha_0 = -10\alpha_1 - 100\alpha_2 - 1000\alpha_3. \quad (5A.5)$$



شكل (١٥-أ)

وهذا يعني أن القيد $b_{10} = 0$ يتضمن قيودا على العلاقة بين α 's، ويكون هذا، أساسا، هو حلنا للنموذج. ولكي نوضح ذلك، دعنا نعود إلى (5A.3) ونعوض عن α_0 من المعادلة (5A.5)، وهذا يعطينا:

$$Y_t = a + \alpha_1 Q_{1t} + \alpha_2 Q_{2t} + \alpha_3 Q_{3t} + u_t, \quad (5A.6)$$

حيث:

$$Q_{1t} = Z_{2t} - 10Z_{1t},$$

$$Q_{2t} = Z_{3t} - 100Z_{1t},$$

وأيضا:

$$Q_{3t} = Z_{4t} - 1000Z_{1t}.$$

وتعد المعادلة (5A.6) الشكل النمطي، وهكذا، فإنه يمكن تقدير كل من $\hat{\alpha}_1$ ، $\hat{\alpha}_2$ و $\hat{\alpha}_3$ باستخدام طرق تقدير الانحدار المتعدد النمطية، افترض أن $\hat{\alpha}_1$ ، $\hat{\alpha}_2$ و $\hat{\alpha}_3$ هي

مقدراتنا، لذا يكون مقدرننا لـ α_0 من المعادلة (5A.5) هو:

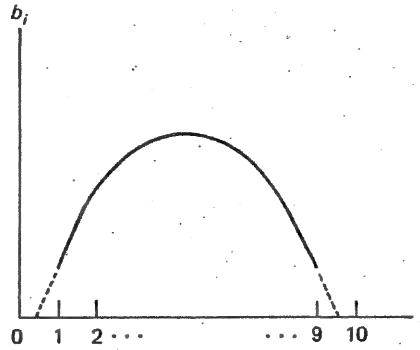
$$\alpha_0 = -10\hat{\alpha}_1 - 100\hat{\alpha}_2 - 1000\hat{\alpha}_3. \quad (5A.7)$$

وأخيرا يمكن اشتقاق مقدراتنا b's من المعادلة (5A.2) على النحو التالي:

$$\begin{aligned} \hat{b}_0 &= \hat{\alpha}_0 \\ \hat{b}_1 &= \hat{\alpha}_0 + \hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3 \\ &\vdots \\ \hat{b}_9 &= \hat{\alpha}_0 + 9\hat{\alpha}_1 + 81\hat{\alpha}_2 + 729\hat{\alpha}_3 \\ \hat{b}_{10} &= 0. \end{aligned} \quad (5A.8)$$

وباختصار، فإن فرض القيد الطرفي $b_{10} = 0$ قد مكنتنا من الإحلال محل α_0 ومن ثم، إسقاط هذا المعامل من نموذج ألون للانحدار (5A.3). ونترك للقارئ على سبيل التدريب، أن يثبت أن فرض القيد $b_0 = 0$ و $b_{10} = 0$ سيؤدي بنا إلى إحلال تعبيرات في α_2, α_3 محل كلا من α_0 و α_1 في نموذج الانحدار المتعدد (5A.3)، وهكذا فإن كل من α_0 و α_1 سوف يختفي من المعادلة التي نقدرها فعلا.

وقد ذكرنا في متن هذا الفصل أن استخدام هذه الطريقة غير المباشرة في فرض قيود طرفية سوف يعطي مقدرات غير متحيزة وذات تباينات أصغر من تلك الناجمة عن استخدام الطريقة المباشرة في التقدير (والتي تتضمن إسقاط كل من $X_{i,k}$ و $X_{i,k}$ من التحليل). ويكون ذلك صحيحا فقط في ظل افتراض محدد، وهو أن المعلومات الطرفية تقع على نفس المنحنى لمتعدد الحدود، كما هو الحال بالنسبة للمعلومات الأخرى غير الصفريّة. فعلى سبيل المثال، إذا كانت القيود الطرفية هي $b_0 = b_{10} = 0$ فإنه ينبغي علينا أن نفترض أن قيمة متعدد الحدود تساوي الصفر عندما تكون $i=0$ ، $i=10$ وهذا ما فعلناه بالنسبة لـ $b_{10} = 10$ في مثالنا السابق (5A.4). ولكن قد لا يتحقق هذا الافتراض، عند التطبيق. افترض، على سبيل المثال، أن $k=10$ وأن $b_0 = b_{10} = 0$ ، افترض أيضا، أن المعلومات غير الصفريّة b_1, \dots, b_a تقع جميعها على متعدد للحدود من الدرجة الثانية مثل الموجود في الشكل رقم (أ٥-٢).



شكل (٢-١٥)

في هذه الحال، يتضمن افتراض أن قيمة متعدد الحدود تساوي الصفر لكل من $i=0$ ، $i=10$ أنه، إذا تخيلنا متعدد الحدود سوف يقطع المحور الأفقي عند $i=0$ و $i=10$. ويظهر شكل (٢-١٥) أن ذلك ليس صحيحا بالضرورة. حيث يظهر الشكل رقم ٢١٥ متعدد الحدود يقطع المحور الأفقي (كما يظهر في الشكل الخطوط المتقطعة) في مكان ما بين $i=0$ و $i=1$ وأيضا بين $i=9$ و $i=10$ ، وهكذا، إذا ما استخدمنا الطريقة السابقة في التقدير (5A.4) فإننا سوف ننتهي بفرض مجموعة من القيود الطرفية غير الصحيحة. وتكون النتيجة هي أن المقدرات الناتجة متحيزة. وباختصار لا ينبغي علينا أن نفرض قيودا طرفية إلا إذا تبين من التحليل المتعمق صحتها. ولهذا السبب فإن الطريقة المباشرة في معالجة المعلومات المرتبطة بقيم العلامات الطرفية (عن طريق، إسقاط X_t مثلا، من (5A.1) إذا كان يعتقد بأن معاملها b_0 يساوي الصفر) تبدو أكثر نجاحا.

ملحق ب (B):

اختبار الفرضيات المتضمنة لأكثر من معلمة انحدار

افترض أنه يوجد نموذج الانحدار التالي:

$$Y_t = a_0 + a_1 X_{1t} + \dots + a_k X_{kt} + u_t, \quad t = 1, \dots, n, \quad (5B.1)$$

حيث نفترض أن المتغيرات المستقلة X_1, \dots, X_k ، وأيضا، الأخطاء، العشوائية تحقق

جميع الفروض التقليدية لنموذج الانحدار، ومنها، بالطبع، افتراض أن الأخطاء العشوائية تكون موزعة توزيعاً طبيعياً.

يرغب الاقتصاديون، عادة، في اختبار فرضيات النماذج المشابهة للنموذج (5B.1) التي تتضمن أكثر من معلمة واحدة، وتأتي هذه الفرضيات في أحد الشكلين التاليين: الأول منهما يرتبط بالقيود الخطية على المعاملات في (5B.1) وأحد الأمثلة لهذه الفرضيات هو:

$$\begin{aligned} a_1 &= a_2, \\ a_3 &= 2a_5, \\ a_0 + a_1 + a_2 + \dots + a_k &= 1. \end{aligned} \quad (5B.2)$$

ويرتبط الشكل الثاني منهما بمعنوية مجموعة من المتغيرات المستقلة، افترض على سبيل المثال، أننا نريد اختبار الفرضية بأن Y_i لا تعتمد على أي من X_1, X_2 أو X_3 في (5B.1) حيثئذ، نكون مهتمين باختبار الفرضية:

$$a_1 = a_2 = a_3 = 0. \quad (5B.3)$$

وفي كلتا هاتين الحالتين السابقتين، أخذت الفرضيات الموضوعة في (5B.2) و (5B.3) على أنها فرضيات العدم، أما الفرضيات البديلة فقد أخذت على أنها مكملات لفرضية العدم ($\text{not } H_0$). على سبيل المثال، فإن الفرضية البديلة لفرضية العدم الموجودة في (5B.3) فرضية وجود واحد، في الأقل، من المعلمات a_1, a_2, a_3 لا يساوي الصفر. تكون الفرضية البديلة للفرضية الموجودة في (5B.2) هو:

$$\begin{aligned} a_1 &\neq a_2, \\ a_3 &\neq 2a_5, \\ a_0 + a_1 + a_2 + \dots + a_k &\neq 1. \end{aligned} \quad (5B.4)$$

ويوجد، لحسن الحظ، منهج مباشر لاختبار مثل هذه الفرضية، ويمكن توضيح هذا المنهج في الخطوات التالية:

الخطوة الأولى: ندخل فرضية العدم موضع الاهتمام في نموذج الانحدار،

مثلاً، إذا كانت الفرضية هي أن $a_1 = a_2$ ، فإننا سوف نعيد كتابة نموذج الانحدار الموجود في (5B.1) على النحو التالي:

$$Y_t = a_0 + a_1(X_{1t} + X_{2t}) + a_3X_{3t} + \dots + a_kX_{kt} + u_t. \quad (5B.5)$$

وعلى سبيل المثال، آخر، إذا كانت فرضية العدم هي $a_1 = a_2 = a_3 = 0$ ، فإننا نعيد كتابة النموذج (5B.1) على النحو التالي:

$$Y_t = a_0 + a_4X_{4t} + \dots + a_kX_{kt} + u_t. \quad (5B.6)$$

وأخيراً، وعلى سبيل توضيح ثالث، إذا كانت فرضية العدم $a_1 + 2a_2 + 5a_3 = 10$ ، فإنه سيكون لدينا $a_1 = 10 - 2a_2 - 5a_3$ ، ولذا يصبح النموذج:

$$Y_t = a_0 + 10X_{1t} + a_2(X_{2t} - 2X_{1t}) + a_3(X_{3t} - 5X_{1t}) + a_4X_{4t} + \dots + a_kX_{kt} + u_t. \quad (5B.7)$$

والذي يمكننا إعادة كتابته على النحو التالي:

$$(Y_t - 10X_{1t}) = a_0 + a_2(X_{2t} - 2X_{1t}) + a_3(X_{3t} - 5X_{1t}) + a_4X_{4t} + \dots + a_kX_{kt} + u_t. \quad (5B.8)$$

الخطوة الثانية: تقدير الشكل المقيد من نموذج الانحدار السابق وحساب

مجموع مربعات الخطأ ESS_R *

الخطوة الثالثة: تقدير الشكل الأولي (غير المقيد) لنموذج الانحدار (5B.1)

وحساب مجموع مربعات الخطأ ESS_U .

الخطوة الرابعة: تحديد الفرق في عدد الملمات بين نموذجي الانحدار المقيد

وغير المقيد، فإذا رمزنا لهذا الفرق بأنه d ، تكون الفرضية المناظرة لـ $a_1 = a_2$ ، على سبيل المثال، في (5B.2) هي $d=1$. وعلى سبيل مثال، آخر نجد أن الفرضية في

(5B.3) تتضمن أن $d=3$.

الخطوة الخامسة: نحسب النسبة التالية:

$$ESS = \sum (Y_t - \hat{Y}_t)^2$$

تذكر أننا عرفنا من الفصلين الثاني والرابع أن

$$\frac{(ESS_R - ESS_U) / d}{ESS_U / (n - k - 1)}, \quad (5B.9)$$

حيث إن n هو عدد المشاهدات و $(k+1)$ هو عدد الملمات في النموذج الأصلي (غير المقيد)، ونقبل فرضية العدم أو نرفضها على أساس حجم هذه النسبة الموجودة في (5B.9) افترض، على سبيل المثال، أن فرضية العدم موضع الاهتمام غير صحيحة. حينئذ، سيكون نموذج الانحدار المقيد مبنيًا على افتراض خاطئ ومن ثم، لن يكون محددًا تحديدًا دقيقًا. ويكون النموذج الأصلي، في هذه الحالة، صحيحًا. ويوضح ذلك أن كلا من ESS_U و ESS_R سيختلفان عن بعضهما بعضًا، وفي حالتنا هذه، نتوقع بتحديد أكثر، أن يكون $ESS_U < ESS_R$.

ومن ناحية أخرى، افترض أن فرضية العدم صحيحة، حينئذ، سيكون نموذج الانحدار المقيد محددًا تحديدًا صحيحًا. ولكن نموذج الانحدار غير المقيد سيكون، بدوره محددًا بطريقة صحيحة أيضًا. وعلى الرغم من أن ذلك قد يبدو غريبًا، إلا أنه ليس من الصعب أن ندرك السبب. ذلك أنه عند التحديد الكامل للنموذج يكون الافتراض الوحيد الذي نختبره هو أن ملمات النموذج ثوابت. ومن الواضح أن هذا الافتراض سوف يتحقق سواء كانت فرضية العدم في (5B.2) أو (5B.3) صحيحة أم لا. ذلك أن شرط $a_1 = a_2$ لا يخالف أي من الافتراضات المرتبطة بنموذج الانحدار كما في (5B.1). و، بالمثل، ولأن الصفر هو مقدار ثابت، فإنه إذا كانت فرضية العدم هي $(a_2 = a_3 = 0)$ فإن هذا، مرة أخرى، لن يؤدي إلى مخالفة أي من الافتراضات في (5B.1).

وفي فصل لاحق من هذا الكتاب، سوف نرى أنه إذا كانت الفرضية موضع الاهتمام صحيحة، فإن تطبيق نموذج الانحدار المقيد سوف يحقق فوائد، ترتبط بخصائص المقدرات. ولكننا، في المرحلة الحالية نحتاج أن نلاحظ، فقط، أنه إذا كان كل من الأشكال المقيدة وغير المقيدة في نموذج الانحدار قد حدد تحديدًا صحيحًا فإننا سوف نتوقع أن يكون الفرق بين ESS_U و ESS_R صغيرًا. وكل ذلك يمكن تلخيصه عن طريق ملاحظة أن القيم الكبيرة للنسبة الموجودة في (5B.9) توحي بأن

فرضية العدم تكون غير صحيحة بينما توحى القيم الصغيرة لها بصحة فرضية العدم.

وباستخدام اللغة الإحصائية يمكن أن نبين* أنه، إذا أعتبر فرضية العدم صحيحة فإن النسبة (5B.9) تأخذ شكل متغير F مع درجات حرية d و $(n-k-1)$ أو باختصار $F_{d,n-k-1}$.

الخطوة السادسة: لما كانت القيم الصغيرة من هذه النسبة مرتبطة بقبول فرضية العدم فإننا نتوقع قبول H_0 عند مستوى معنوية 5% إذا كانت:

$$\frac{(ESS_R - ESS_U) / d}{ESS_U / (n - k - 1)} < F_{d,n-k-1}^{0.95}, \quad (5B.10)$$

حيث إن $F_{d,n-k-1}$ تكون:

$$\text{Prob}(F_{d,n-k-1} < F_{d,n-k-1}^{0.95}) = 0.95. \quad (5B.11)$$

ويمكننا الحصول على $F_{d,n-k-1}$ من أي جدول من جداول توزيع F . ويوجد أحد هذه الجداول في نهاية الكتاب [الجدول الإحصائي رقم (٣)].

ولتوضيح الخطوات السابقة، دعنا نأخذ مثالا، افترض أن لدينا الافتراض التالي $a_1 = a_2 = a_3 = 0$. افترض أيضا، أن $n=49$ و $k=8$ ، نجد من الجدول الإحصائي رقم (3) أن: $F_{3,40}^{0.95} = 2.84$ ، افترض أننا حددنا ESS_R وأن $ESS_U = 20$ ، تكون نسبتنا حيث هي:

$$\frac{(35 - 20) / 3}{20 / 40} = \frac{15 / 3}{1 / 2} = 10 > 2.84.$$

وهذا يعني أننا سنرفض عند مستوى معنوية 5% الفرضية: $a_1 = a_2 = a_3 = 0$.

* يرجع إلى: Arther Goldberger. *Econometric Theory* (New York, 1964), pp. 173-177.

أسئلة

- ١- افترض أن دالة للإنتاج تأخذ الشكل التالي: $Q_t = (1/A) L_t^a K_t^b e^{u_t}$ ، حيث Q_t ، L_t و K_t هي الإنتاج والعمل ورأس المال في الزمن t على التوالي، وأن u_t هي الأخطاء العشوائية المناظرة. افترض، أيضا، أن $E(u_t) = 0$ ، $E(u_t^2) = \sigma^2$ و u_t مستقلة عن L_t و K_t . اقترح إحدى الطرق لتقدير a ، b و A .
- ٢- افترض أن حجم الاستثمار الخاص في سنة معينة يعتمد على معدل الفائدة وعلى الحزب السياسي الذي ينتمي إليه الرئيس، بمعنى أن الاستثمار يكون أكثر ارتفاعا في حالة ما إذا كان الرئيس ينتمي إلى الحزب الجمهوري بدلا من الحزب الديمقراطي. كون نموذجا يأخذ البيانات في شكل سلسلة زمنية على افتراض أن هناك حزبين فقط.
- ٣- افترض النموذج التالي للانحدار:

$$C_t = a_0 + a_1 F_t Y_t + a_2 Y_t^{1/2} + a_3 (1/A_t) + u_t,$$

حيث C_t : الإنفاق الاستهلاكي للأسرة t ، Y_t دخل تلك الأسرة، F_t حجم الأسرة و A_t حجم الأصول السائلة التي تمتلكها الأسرة، حول هذا النموذج إلى نموذج خطي.

- ٤- افرض أنك؛ تريد تقدير دالة الاستهلاك الخطية البسيطة التالية $C_t = a + bY_t + u_t$ لعدد n من الأفراد. كيف يمكنك أن تأخذ في الحسبان الانتقال في الدالة بين المستهلكين في الحضر والمستهلكين في الريف، إذا كان الحد الثابت من الدالة يتأثر بموقع الإقامة للفرد.
- ٥- افترض أن الإنفاق الاستثماري لإحدى المنشآت يعتمد على معدل الفائدة ومعدل الأرباح وأخيرا على معدل التغير في المبيعات باعتباره مؤشرا على التوقعات:

(أ) كون نموذج الانحدار المقابل.

(ب) افترض أنه، خلال فترة العينة، كانت أرباح هذه المنشأة تعادل خلال

مختلف الفترات الزمنية ١٥٪. ناقش مشاكل التقدير التي تنتج عن ذلك.

٦- افترض أن معدل الاستثمار في فترة معينة t ، I_t يعتمد على معدل الفائدة في تلك الفترة r_t ، وعلى المبيعات في تلك الفترة، وسبع قيم ذات فترات إبطاء لمعدل المبيعات S_t ، افترض، أيضا، أن الأوزان المناظرة لفترات الإبطاء هذه تتزايد في البداية حتى تصل إلى ذروتها ثم تتناقص بعد ذلك.

(أ) كون شكلا غير مقيد لهذا النموذج

(ب) كون شكل آلمون لهذا النموذج

(ج) اكتب المعادلات الطبيعية لشكل آلمون السابق.

٧- افترض أنه يوجد لدينا نموذج الانحدار المتعدد التالي:

$$Y_t = a + b_0 X_t + \dots + b_6 X_{t-6} + \varepsilon_t.$$

افترض أننا استخدمنا منهج آلمون مع متعدد حدود من الدرجة الرابعة لتقدير معالم هذا النموذج. افترض أخيرا أن النتائج كانت، على النحو التالي:

$$\hat{\alpha}_1 = 3, \hat{\alpha}_2 = 5, \hat{\alpha}_3 = 4, \hat{\alpha}_4 = -10.$$

(أ) ماذا يكون تقديرنا لـ b_2 ؟

(ب) افترض أننا جعلنا $b_0 + b_1 + \dots + b_6 = 1$. عير عن هذه المعلومة في شكل قيد على α 's في تقريب متعدد الحدود لآلمون.

٨- يقال، أحيانا، إن النتائج التطبيقية لنموذج الانحدار يجب ألا تستخدم للتعنبؤ بالحوادث التي تقع بعيدا جدا عن مجال التجربة. ناقش هذه العبارة (مساعدة للحل: راجع الافتراضات المرتبطة بتقريب متعددات الحدود).

٩- حول نموذج كويك التالي إلى شكل أبسط:

$$Y_t = u_0 + a_1 X_t + b_0 Z_t + b_1 Z_{t-1} + \dots + u_t,$$

$$b_i = b_0 \lambda^i, \quad i=1,2,\dots$$

حيث

١٠- اعتبر نموذج الإبطاء التالي لآلمون:

$$Y_t = b + b_0 X_t + b_1 X_{t-1} + \dots + b_{10} X_{t-10} + u_t$$

حيث نفترض التقريب التالي من الدرجة الثانية:

$$b_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2, \quad i = 0, 1, \dots, 10.$$

افترض أن $b_5 = 3$. المطلوب هو أن تشتق القيود المناظرة على α 's، ومن ثم، الشكل المقيد لنموذج الانحدار.

١١ - افترض أن الإنفاق الاستهلاكي C_t يعتمد على الدخل Y_t وعلى معدل الفائدة r_t إذا كانت r_t تزيد عن 0.05، أما إذا كانت r_t لا تزيد عن 0.05 فإن C_t تعتمد فقط على Y_t . كون هذه العلاقة في شكل نموذج للانحدار.

١٢ - حول النموذج التالي إلى نموذج انحدار خطي متعدد، ثم اكتب المعادلات الطبيعية له:

$$\log Y_t = a_0 + a_1 e^{x_{1t}} + a_2 \left(\frac{1}{1 + X_{1t} X_{2t}} \right) + u_t.$$

مشاكل في تحليل الانحدار

تناولنا في الفصول السابقة طرق تدير العلاقة بين مجموعة من المتغيرات، وتعلمنا كيف يمكننا استخدام هذه المعلومات في اختبار الفرضيات وفي توقع تأثير تغير أحد المتغيرات على الآخر. تعتمد طرق التقدير هذه على عدد من الفروض التي ناقشناها في معالجتنا لنموذج الانحدار. ولكن، عند التطبيق، يجد الاقتصاديون أن من المؤكد (أو المرجح في الأقل في بعض الحالات الأخرى) أن واحداً أو أكثر من هذه الفروض لن يتحقق. إما بسبب طبيعة العلاقة الدالية التي تجمع بين متغيرات النموذج، أو نتيجة لوجود صعوبات كبيرة تنشأ من مجموعة معينة من القيم المشاهدة للمتغيرات. وتوجه الآن معظم الأعمال الحديثة في مجال الاقتصاد القياسي نحو تطوير طرق تقدير معدلة وبنائها لتعالج مثل هذه المشكلات. وسيكون ذلك هو موضوع الفصلين الأخيرين من هذا الكتاب. سنتعرض هنا للحالات التي تسبب مخالفة افتراضات نموذج الانحدار، أو التي تسبب - في الأقل - في إيجاد صعوبات تمنع الاستفادة منها أو تقلل فاعليتها، وبعد ذلك سنناقش الكيفية التي يمكن أن نعدل بها طرق تقديرنا لنأخذ في الاعتبار هذه المشاكل.

(٦-١) تعدد العلاقات الخطية

ناقشنا، بالفعل، في أماكن عديدة، مشكلة تعدد العلاقات الخطية multicollinearity وبلغة فنية تنشأ هذه المشكلة عندما يكون واحد، في الأقل، من المتغيرات المستقلة توليفة خطية من المتغيرات الأخرى. وينتج عن ذلك وجود عدد قليل جدا من المعادلات الطبيعية المستقلة، ومن ثم، عدم امكانية اشتقاق مقدرات للمعاملات الموجودة بالنموذج كافة. وعلى سبيل مراجعة مختصرة افترض أن:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t. \quad (6.1)$$

حيث تكون قيم كل من X_1 و X_2 دائما منطبقة على بعضها بعضا بمعنى أن لدينا:

$$X_{1t} = X_{2t}, \quad t=1,2,\dots,n.$$

وهذا يعني أن أي تحرك من فترة لأخرى في المتغير X_1 يناظر، تماماً، تحرك مماثل في X_2 . فإذا كان ذلك صحيحاً فإنه لا يمكننا أن ن عزل تأثير X_1 على Y عن تأثير X_2 على Y . لكن، تذكر أنه مازال من الممكن تقدير الأثر المشترك لكل من X_1 و X_2 على Y أي أنه، بإحلال X_{1t} محل X_{2t} في المعادلة (6.1)، نستطيع تقدير قيمة b_3 حيث إنها تساوي $(b_1 + b_2)$ في المعادلة (6.2)*.

$$Y_t = b_0 + (b_1 + b_2) X_{1t} + u_t = b_0 + b_3 X_{1t} + u_t. \quad (6.2)$$

وإذا ما استمر وجود تعدد العلاقات الخطية بين X_{1t} و X_{2t} في فترات خارج الفترة التي تغطيها عينتنا فإن الصيغة المقدرة للمعادلة (6.2) تصبح ملائمة للتنبؤ بالقيم المستقبلية للمتغير التابع Y_t .

ويطلق على حالة تعدد العلاقات الخطية هذه تعدد العلاقات الخطية التام، حيث يكون واحداً أو أكثر من المتغيرات المستقلة توليفة خطية تامة من المتغيرات الأخرى. وتنتج هذه الحالة عن الطريقة التي تكون بها معادلة الانحدار، وتشبه هذه الحالة تلك التي ناقشناها في الفصل السابق حيث استخدمت المتغيرات الصورية

* يمكننا الآن رياضياً أن نقدر المعادلة (6.2) ذلك لأننا قللنا عدد العلامات التي ينبغي أن نقدرها بمقدار معلمة واحدة، ولذلك، فإن عدد العلامات أصبح يساوي عدد المعادلات الطبيعية المستقلة.

لكل الفصول المناخية الأربعة (بدلاً من ثلاثة فقط). وعلى سبيل مثال آخر، افترض أننا وضعنا دالة الاستهلاك في الشكل:

$$C_t = b_0 + b_1 X_{dt} + b_2 Y_{d(t-1)} + b_3 (\Delta Y_{dt}) + u_t, \quad (6.3)$$

حيث Y_{dt} تعكس تأثير الدخل الجاري على الاستهلاك $Y_{d(t-1)}$ تشير إلى تأثير مستويات الدخل الماضية أو تأثير العادة، بينما تعكس ΔY_{dt} و $[\Delta Y_{dt} = Y_{dt} - Y_{d(t-1)}]$ ، أثر التوقع الناجم عن التغيرات في مستويات الدخل، ولما كانت ΔY_{dt} توليفة خطية من كل من $Y_{d(t-1)}$ و Y_{dt} فلن يكون باستطاعتنا تقدير كل المعاملات في المعادلة (6.3). وسوف نترك للقارئ أن يضع المعادلة (6.3) في صيغة يمكن تقديرها واستخدامها في التنبؤ بقيم C_t .

هذه إذن حالات من تعدد العلاقات الخطية التام. ولكن هذه المشكلة تظهر أيضاً، بدرجات مختلفة، وفي العادة تأتي بدرجة أقل من تعدد العلاقات الخطية التام، وهذه الأخيرة هي التي تسبب معظم المشاكل للباحثين. وتنشأ هذه المشكلة عندما ترتبط المتغيرات المستقلة ببعضها ارتباطاً قوياً، ولكن غير تام. افترض على سبيل المثال أننا نحاول تقدير دالة الطلب كما في المعادلة (6.4).

$$Q_t = b_0 + b_1 P_t + b_2 Y_t + u_t. \quad (6.4)$$

لمجموعة معينة من السلع، ولتكن الواردات، حيث نفترض أن الكمية المطلوبة (Q) تعتمد على مستوى الأسعار للسلع المنتجة محلياً (p) وعلى مستوى دخل المستهلكين. ومن المعروف أنه إذا كانت الأسعار المحلية مرتفعة زاد الطلب على المنتجات الأجنبية التي تعد أرخص نسبياً، ولذا، نتوقع أن تكون b_1 موجبة. وبالمثل، فإنه إذا كانت دخول المستهلكين أكبر كان الطلب أكبر على السلع (بما فيها الواردات)، ولذا نتوقع أن تكون b_2 موجبة أيضاً. ونلاحظ بفحص البيانات المتاحة، أنه على مدى الفترات الزمنية التي يزداد فيها معدل التضخم المحلي، تزداد الواردات، وعلى نحو مشابه، تنمو الواردات أيضاً، خلال فترات تزايد الدخل. والصعوبة هنا هي أن فترات تزايد الدخل هي، عموماً، فترات التضخم المرتفع والعكس صحيح. أو بمعنى آخر، يوجد ارتباط موجب قوي (وإن لم يكن تاماً) بين p و Y .

هذا الارتباط القوي يجعل من الصعوبة بمكان عزل آثار p و Y على Q فعندما يكون هناك تزايد سريع في الواردات في الوقت نفسه الذي يزداد فيه كل من الدخل والأسعار المحلية، يكون من الصعب تحديد الآثار النسبية لكل من التضخم والدخول الأعلى في حفز الزيادة في الواردات.

تعدد العلاقات الخطية غير التام: بعض النتائج المنطقية

كيف يعرف الباحث أن لديه مشكلة تعدد علاقات خطية خطيرة في نموذج؟ كما أشرنا من قبل تظهر مشكلة تعدد العلاقات الخطية في درجات مختلفة وقد تسبب مشاكل عسيرة في بعض الحالات وقد لا تكون كذلك في حالات أخرى. وهناك على أي حال، بعض نتائج الانحدار التي يمكن تفسيرها، فقط، بدلالة الدرجات العالية من تعدد العلاقات الخطية غير التام: منها مثلاً عندما يوجد معامل تحديد كبير R^2 مصاحب لتقديرات غير معنوية إحصائياً لمعاملات المتغيرات المستقلة. ويعني ذلك أن هناك متغيرات مستقلة معينة (أو في الأقل، واحد منها) تؤثر تأثيراً منتظماً في المتغير التابع (كما يشار إليه بوساطة قيمة R^2 المرتفعة) ولكن لا يمكننا معرفة أي منها هو المسئول عن ذلك التأثير.

وتظهر مشكلة تعدد العلاقات الخطية بدقة أكثر في وجود تباينات كبيرة لمقدرات المعاملات. ولما كان التباين الكبير يعني أن نسبة مئوية (مثلاً 95%) لفترة ثقة معينة للمعلمة المناظرة ستكون عريضة نسبياً، فإن مدى كبير من القيم للمعلمة (وربما تتضمن هذه قيمة الصفر) ستكون متسقة مع فترتنا للثقة. ويعني ذلك - أنه حتى إذا كان المتغير المستقل المناظر له تأثير مهم على المتغير التابع، فإن مشكلة تعدد العلاقات الخطية قد تجعل من الصعب القيام بتقدير أثر ذلك المتغير بدقة. ومن ثم، تنخفض درجة الثقة في السياسات المقترحة بناء على هذه المقدرات.

وحتى نرى كيف تنتج تباينات كبيرة عن مشكلة تعدد العلاقات الخطية غير التام (ومن ثم، انحرافات معيارية كبيرة) لمقدراتنا، تذكر أننا في الفصل الرابع أن تباين المقدّر \hat{b}_1 هو:

$$\text{var}(\hat{b}_i) = \frac{\sigma_u^2}{\sum \hat{v}_{ii}^2} \quad (6.5)$$

حيث \hat{v}_{ii} هو الباقي في انحدار المتغير المستقل رقم i^{th} (X_{ii}) على جميع المتغيرات المستقلة الأخرى في النموذج $\hat{v}_{ii} = X_{ii} - \hat{X}_{ii}$ فإذا كانت هناك علاقة خطية وثيقة بين المتغيرات المستقلة، فإن \hat{v}_{ii} سيصبح صغيراً لأن القيمة المحسوبة \hat{X}_{ii} سوف تكون قريبة من القيمة الفعلية للمتغير X_{ii} . ويتبع عن ذلك صغر المقام في المعادلة (6.5) مما يوجد تبايناً كبيراً لـ \hat{b}_i .

ومن المهم أن نفسر بدقة مايعنيه كل هذا. فيجب أن نتذكر أن ذلك لايعني أن تقديرات المعاملات متحيزة، بل أنها تظل غير متحيزة، كما أثبتنا ذلك في ملحق الفصل الرابع. غير أن مايؤدي إليه تعدد العلاقات الخطية هو عدم دقة مقدراتنا: حيث يصبح تباينها كبيراً، ومن ثم، لايمكن الاعتماد عليها اعتماداً كبيراً. وتكمن المشكلة - كما ذكرنا من قبل - في صعوبة تحديد تأثير كل متغير مستقل - بعيداً عن تأثير المتغيرات الأخرى على المتغير التابع.

تعليق إضافي

افترضنا، ضمناً حتى الآن، أن جميع المتغيرات المستقلة بالنموذج مرتبطة ببعضها البعض ارتباطاً قوياً. وقد لا يكون الوضع كذلك بالضرورة. افترض مثلاً، أن نموذج الانحدار يحتوي على متغيرات مستقلة X_{1t} , X_{2t} , X_{3t} . افترض، أيضاً، أن X_{1t} و X_{2t} مرتبطتان ببعضهما بعضاً ارتباطاً قوياً (وإن كان هذا الارتباط ليس تاماً) ولكن X_{3t} ليست مرتبطة نسبياً بكل من X_{1t} و X_{2t} . حينئذ فإن صيغة التباين في المعادلة (6.5) تفيد أن تباين المقدرات المناظرة للمعاملات X_{1t} و X_{2t} (ولتكن \hat{a}_1 و \hat{a}_2) ستكون كبيرة بينما يكون تباين $(\hat{a}_3)X_{3t}$ وهو معامل X_{3t} ليس كذلك بالضرورة. وبديهيًا إذا كانت X_{3t} ليست مرتبطة بقوة بكل من X_{1t} و X_{2t} فإن \hat{v}_{3t}^2 يمكن أن تكون كبيرة، عمومًا طالما أن \hat{X}_{3t} تمثل تنبؤاً غير دقيق لـ X_{3t} (بمعنى أن X_{3t} لا يمكن تفسيرها مرضياً بوساطة X_{1t} و X_{2t}).

والآن يمكننا أن نضع قاعدة قد تكون مفيدة في تفسير مشكلة تعدد العلاقات الخطية. ذلك أن المقدار $\sum \hat{v}_{it}^2$ والذي يظهر في مقام المعادلة (6.5) هو مجموع مربعات الخطأ الذي حصل عليه عن طريق تكوين انحدار لـ X_{it} على جميع المتغيرات المستقلة الأخرى الموجودة بالنموذج. دعنا نرمز إلى مجموع مربعات الخطأ هذا بالرمز ESS_i وأن $TSS_i = \sum X_{it}^2$ وحيث $TSS_i = RSS_i + ESS_i$ ومنها فإن $RSS_i = TSS_i - ESS_i$. ومن الواضح أن $ESS_i \geq 0$ وأن $TSS_i \geq 0$. ويمكن أن نثبت، أيضا، أن $RSS_i \geq 0$ ، ومن ثم، فإن $TSS_i \geq RSS_i$ و $TSS_i \geq ESS_i$.

دعنا نرمز للنسبة RSS_i/TSS_i بالرمز r_i^2 . حيث، فإن $0 \leq r_i^2 \leq 1$. ومن الواضح أنه كلما كانت قيم X_{it} مرتبطة ارتباطا قويا بقيم المتغيرات المستقلة الأخرى كانت ESS_i صغيرا، ومن ثم، تكون r_i^2 قريبة من الواحد الصحيح. وعلى العكس من ذلك، كلما كانت العلاقة بين X_{it} والمتغيرات المستقلة الأخرى ضعيفة، كانت ESS_i كبيرة، ومن ثم، تكون r_i^2 قريبة من الصفر. لذا يكون r_i^2 مشابه لمعامل التحديد المتعدد الذي يربط X_{it} بجميع المتغيرات المستقلة الأخرى في النموذج.

والآن دعنا نعبر عن مقام صيغة التباين بالمعادلة (6.5) صورة أخرى من خلال معرفة $ESS_i = TSS_i - RSS_i \equiv TSS_i (1 - r_i^2)$ ويتبع عن ذلك أن تباين \hat{b}_i يمكن أن نعبر عنه على النحو التالي:

$$\text{var}(\hat{b}_i) = \frac{\sigma_u^2}{TSS_i(1-r_i^2)} \equiv \frac{\sigma_u^2}{\sum X_{it}^2(1-r_i^2)} \quad (6.6)$$

ويشير التعبير الموجود في المعادلة (6.6) بوضوح لتأثير تعدد العلاقات الخطية الجزئي على تباين أحد المقدرات. وينظر غياب تعدد العلاقات الخطية الجزئي الحالة التي يكون فيها $r^2 = 0$. وتصبح المشكلة أكثر صعوبة وتؤدي إلى تباين أكبر كلما اقتربت r من الواحد الصحيح. لاحظ أخيرا أن r ليس من الضروري أن يتساوى في الكبر لكل من $k, \dots, i=1$.

* أنظر (4A.5) في الملحق للفصل الرابع ولاحظ أنه، بالنسبة للمتغير المستقل رقم k^{th} يكون $TSS_k' = \sum X_{kt}^2$

و $ESS_k = \sum \hat{v}_{kt}^2$ وأن $RSS = \sum \hat{v}_{kt}^2$ وذلك لأن $\sum \hat{X}_{kt} \hat{v}_{kt} = 0$

بعض الحلول

ليس من السهل حل مشكلة تعدد العلاقات الخطية، غير أنه يمكن للباحث، دائماً، أن يحاول زيادة دقة مقدراته (أي تخفيض تبايناتها) عن طريق زيادة عدد المشاهدات. وعلى سبيل المثال، فإنه، مهما صغرت قيم v_{ii}^2 في المعادلة (6.5) فإنه من الواضح أن تباين (\hat{b}_i) يتناقص مع زيادة حجم العينة. ولكن، من الواضح، أيضاً، أن زيادة عدد مشاهدات العينة ليس ممكناً دائماً، وفي حالة كون مشكلة تعدد العلاقات الخطية خطيرة بما فيه الكفاية، فإن زيادة حجم العينة قد لا يساعد كثيراً اللهم إلا إذا كانت هذه الزيادة كبيرة جداً.

ومن الطرق البديلة لعلاج مشكلة تعدد العلاقات الخطية إضافة معلومات يمكن استخدامها في تقدير قيم المعاملات الفردية. افترض، على سبيل المثال، أننا نرغب في تقدير دالة الإنتاج الميينة في المعادلة (6.7) لسلعة ما:

$$Q_t = AL_t^\alpha K_t^\beta e^{u_t}, \quad (6.7)$$

حيث ترمز Q_t الكمية المنتجة في الفترة t ، L_t لساعات العمل، K_t لرأس المال، u_t للخطأ العشوائي و α, β, A هي المعلمات التي ينبغي تقديرها. تذكر أنه يمكن، عن طريق استخدام التحويل اللوغاريتمي وضع المعادلة رقم (6.7) في شكل يمكن تقديره على النحو التالي:

$$Q_t^* = A^* + \alpha L_t^* + \beta K_t^* + u_t. \quad (6.8)$$

حيث ترمز النجوم إلى لوغاريتمات المتغيرات الأولية في المعادلة (6.7). افترض، للتوضيح، أنه توجد لدينا مشكلة تعدد علاقات خطية جزئي في العينة موضع البحث يأخذ شكل ارتباط قوي بين L و K في هذه الحال، يؤدي الارتباط القوي بين L و K (ضمن أشياء أخرى) إلى إيجاد تباينات كبيرة لمقدرات معلمات مرونة دالة الإنتاج α و β .

والآن، دعنا نفترض أنه توجد دلائل قوية، بناء على المعلومات التي أمكن الحصول عليها من مصدر آخر، تشير إلى أن هذه الصناعة المنتجة لهذه السلعة

تخضع لقانون ثبات غلة الحجم. ومن مناقشتنا لدوال الإنتاج في الفصل الأخير، نجد أن هذا يعني أن: $(\beta + \alpha = 1)$ ويمكننا الآن بوساطة هذه المعادلة أن نستبدل β بـ $(1 - \alpha)$ في المعادلة (6.7) لكي نحصل على:

$$Q_t = AL_t^\alpha K_t^{(1-\alpha)} e^{u_t}. \quad (6.9)$$

وبأخذ اللوغاريتمات، يمكننا أن نحصل على:

$$Q_t^* = A^* + \alpha L_t^* + (1 - \alpha) K_t^* + u_t, \quad (6.10)$$

حيث ترمز النجوم - مرة أخرى - إلى لوغاريتمات المتغيرات الأولية. وبإعادة ترتيب مكونات المعادلة (6.10)، نحصل على:

$$Q_t^* - K_t^* = A^* + \alpha(L_t^* - K_t^*) + u_t \quad (6.11)$$

$$Y_t^* = A^* + \alpha Z_t^* + u_t, \quad \text{أو}$$

$$Y_t^* = (Q_t^* - K_t^*) \text{ و } Z_t^* = (L_t^* - K_t^*). \text{ حيث:}$$

وهكذا فإن معلوماتنا المسبقة مكنتنا من اختزال نموذجنا إلى نموذج يحتوي على متغير مستقل واحد Z_t^* . وينبغي أن يلاحظ القارئ أنه حتى ولو كانت L_t^* و K_t^* مرتبطتين ببعضهما بعضا ارتباطا قويا فلن توجد، عموماً، مشكلة ناتجة عن تعدد العلاقات الخطية في تقدير المعادلة (6.11)، فإذا افترض، مثلاً أن $L_t^* = 4K_t^*$ فإنه لن توجد مشكلات تقدير بسبب تعدد العلاقات الخطية نتيجة لأن نموذجنا (6.11) سوف يختزل إلى:

$$Y_t^* = A^* + \alpha(3K_t^*) + u_t. \quad (6.12)$$

وباختصار فإن المعلومات الإضافية التي تبين أن الصناعة تخضع لقانون ثبات غلة الحجم قد مكنتنا من الحصول على مقدرات ذات تباينات أقل للمعلمات A و α .

* يبين هذا المثال أننا سنواجه مشكلة، فقط، إذا كانت $L_t^* - K_t^*$ تساوي مقدار ثابت. ويحدث ذلك عندما تكون L_t نسبة من K_t أي عندما يتحقق $L_t = dK_t$ ، حيث تمثل d مقدار ثابت. وفي هذه الحالة الخاصة ستكون $Z_t^* = (L_t^* - K_t^*)$ في المعادلة (6.11) مقداراً ثابتاً وسوف لاتمكن من الحصول على مقدار لـ α . وفي هذا المجال تذكر من مناقشتنا لنموذج الانحدار البسيط في الفصل الثاني أن المتغير المستقل ينبغي أن يأخذ، في الأقل، قيمتين مختلفتين من أجل أن تتمكن من تقدير معاملات الانحدار.

ومقدرنا لـ β سوف يصبح، حيثئذ:

$$\hat{\beta} = 1 - \hat{\alpha}. \quad (6.13)$$

تأثيره على التنبؤ

لا تتوافر في كثير من الأحيان معلومات إضافية يمكن استخدامها لتخفيف حدة مشكلة تعدد العلاقات الخطية، ويصطدم الباحث بمجموعة من تقديرات المعلومات لا يمكن الاعتماد عليها. وربما نشعر بشئ من الراحة إذا علمنا أن المعادلة المقدرة ماتزال مقبولة لأغراض التنبؤ. وبأخذ مثال شاذ، افترض أن العلاقة التالية التي تظهر تعددا تاما للعلاقات الخطية وذلك بسبب الطريقة التي تعرف بها المتغيرات.

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t, \quad (6.14)$$

حيث $X_{1t} = 3X_{2t}$. كما لوحظ من قبل وبسبب أن X_1 ترتبط بعلاقة خطية تامة مع X_2 ، فمن غير الممكن تقدير أي من b_1 و b_2 . ولكن لغرض التنبؤ لانهتم بقيم b_1 و b_2 في ذاتها ولكننا نهتم بالقيمة المتوسطة لـ Y_t المناظرة لـ X_1 و X_2 أي:

$$\begin{aligned} Y_t^m &= b_0 + b_1 X_{1t} + b_2 X_{2t} \\ &= b_0 + (3b_1 + b_2) X_{2t}. \end{aligned} \quad (6.15)$$

ونعرف من مناقشتنا الحالية أنه يمكننا تقدير القيمة المتوسطة لـ Y_t بوساطة المقدر:

$$\hat{Y}_t = \hat{b}_0 + \hat{b}^* X_{2t}, \quad (6.16)$$

حيث \hat{b}^* هي مقدرنا للمجموع $(3b_1 + b_2)$. وهكذا، فعلى الرغم من أننا لا نستطيع تقدير تأثير كل من X_1 و X_2 على Y_t إلا أنه يمكننا تكوين تنبؤات ترتبط بقيمة Y_t المناظرة لأي قيمة من X_{2t} طالما ظلت العلاقة $X_{1t} = 3X_{2t}$ قائمة. * وتمتد هذه النتيجة لتشمل حالات تعدد العلاقات الخطية غير التام، غير أن إثبات ذلك يخرج عن

* ينبغي أن يكون واضحاً أن دقة التنبؤ أو جودته تعتمد على اعتبارين اثنين هما: (أ) دقة مقدراتنا للقيمة المتوسطة لـ Y_t (أي تباين \hat{Y}_t باعتباره مقدراً لـ Y_t^m)، و (ب) حجم تباين الخطأ العشوائي u_t . ولسوء الحظ، فإن المعادلات التي تصف دقة التنبؤ في إطار الانحدار المتعدد تتطلب معرفة مفاهيم إحصائية فوق المستوى المستخدم في هذا الكتاب.

نطاق هذا الكتاب. على الرغم من أن مقدرتنا للتأثيرات المنفصلة للمتغيرات المستقلة قد تكون لها تباينات كبيرة، فإن مقدرنا \hat{Y}_t للأثر المشترك على Y_t ، الذي يمكن أن يعبر عنه بوساطة القيمة المتوسطة لـ Y_t ، أو Y_t^m ، قد يكون له تباين صغير. ولما كانت تنبؤاتنا تتضمن تقدير متوسط Y_t ، وأنه يمكننا تقدير المتوسط بدقة كبيرة، فإن مشكلة تعدد العلاقات الخطية قد لا تمثل مشكلة صعبة لأغراض التنبؤ.

(٦-٢) مشكلة الارتباط الذاتي

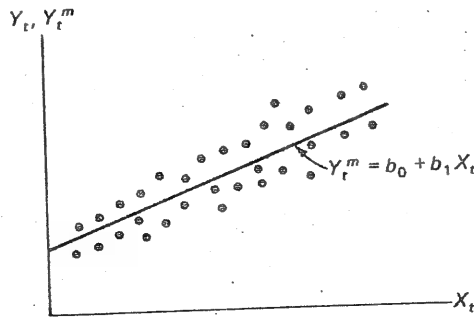
نعلم أن أحد الفروض الأساسية لنموذج الانحدار هو أن قيمة الخطأ العشوائي في إحدى الفترات الزمنية تكون مستقلة عن قيمته في أي فترة زمنية أخرى بحيث يمكن القول إن:

$$\text{COV}(u_t, u_s) = 0, \quad t \neq s.$$

وتتضمن هذه المعادلة بالنسبة لقيم معطاة للمتغيرات المستقلة أن قيمة Y_t سوف تختلف عن قيمتها المتوسطة Y_t^m بمقدار مستقل عن حجم الخطأ العشوائي في أي فترة زمنية أخرى. وبالنسبة لنموذج الانحدار الذي يأخذ الشكل:

$$Y_t = b_0 + b_1 X_t + u_t,$$

نتوقع أن تأخذ خريطة الانتشار الصورة الموضحة في الشكل (٦-١) حيث تكون النقط المشاهدة «متشعبة عشوائيا» حول خط الانحدار.

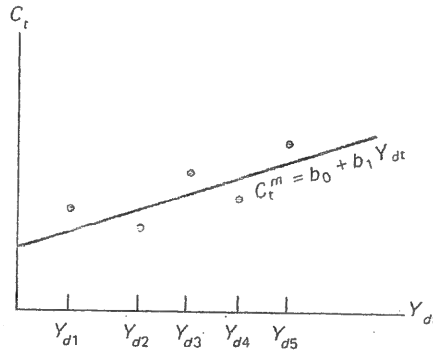


شكل رقم (٦-١)

افترض من ناحية أخرى، أن $\text{cov}(u_t, u_s) \neq 0$ ، لذلك، تكون القيم المتتابة للخطأ العشوائي غير مستقلة عن بعضها بعضا. افترض، على سبيل المثال، أنه يمكن وصف السلوك الاستهلاكي لفرد ما بوساطة المعادلة:

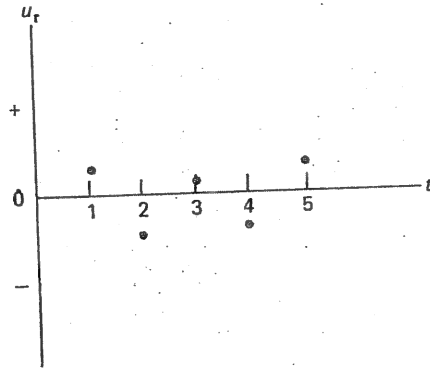
$$C_t = b_0 + b_1 Y_{dt} + u_t, \quad (6.17)$$

ولكن قيم u_t ليست مستقلة عن قيمها السابقة. فإذا كان هذا الفرد ينفق، مثلاً، كثيراً، جداً في الفترة الأولى (ربما نتيجة زيارة غير متوقعة من بعض أصدقائه) بحيث تكون $u_1 > 0$ فسوف يحاول أن يعوض هذه الزيادة في الإنفاق في الفترة التي تليها عن طريق إنفاق قدر أقل من المعتاد، ومن ثم، نتوقع أن تكون $u_2 < 0$. لاحظ أن ذلك يتضمن بعمومية أكثر أن u_t لها ارتباط سالب مع u_{t+1} . فإذا كان مستوى الدخل يزداد في الفترات المتعاقبة فإن مثل هذا التصاحب السالب بين القيم المتتابة للخطأ العشوائي يتوقع أن يولد خريطة انتشار تشبه، إلى حد ما، الشكل رقم (٦-٢).



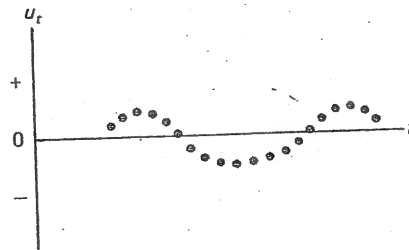
الشكل رقم (٦-٢)

فإذا أردنا أن نرسم شكلاً لقيم الخطأ العشوائي على مدى الزمن، فإننا نتوقع الحصول على شكل يشبه الشكل رقم (٦-٣).



الشكل رقم (٦-٣)

وبالمقابل، قد تظهر قيم الخطأ العشوائي ارتباطاً موجباً عبر الزمن نتيجة (على سبيل المثال) لبطء التكيف في السلوك الاقتصادي، هنا، قد نجد أن القيمة الموجبة لـ u تتبع بقيمة موجبة أخرى لـ u ، وقيمة سالبة تتبعها قيم سالبة أخرى $\text{cov}(u_t, u_{t+1}) > 0$. وعلى افتراض أن قيم الخطأ العشوائي تتحدد بوساطة قيم خارجية (ستحدد بدقة أكثر فيما بعد) تتحول آثارها عشوائياً من الموجب إلى السالب، فإنه يمكننا أن نتوقع نمط قيم u عبر الزمن والذي يحتوي على نتائج مختلفة من قيم موجبة وسالبة. وعلى سبيل المثال إذا كانت القوى الخارجية تولد قيماً موجبة لـ u ، فإنها سوف تتبع بقيم موجبة أخرى، وإذا نتج عن هذه القوى قيم سالبة، فإنه سوف تتبعها قيم سالبة إضافية، وهكذا، يظهر مثل هذا النمط في الشكل رقم (٦-٤).



شكل (٦-٤)

وتعرف مشكلة الاعتماد المتبادل بين القيم المتتابعة للخطأ العشوائي بمشكلة الارتباط الذاتي autocorrelation. وسوف نبين أنه، في ظل افتراضات معينة، بالإضافة لافتراضاتنا السابقة، فإنه، عند أي قيمة معطاة للمشاهدة $E(u_t) = 0$ ، سوف لا يكون الخطأ العشوائي مرتبطا بالمتغيرات المستقلة بحيث تظل مقدرات المعلمات غير متحيزة. وكما سنرى فيما بعد، فإن مشكلة الارتباط الذاتي تثير قلقا بشأن قيمة تباين المقدرات. وعلى نحو خاص، فإن الصيغ التي اشتقت للتباينات تصبح غير صحيحة إذا وجدت مشكلة الارتباط الذاتي. فإذا استمرينا في استخدام هذه الصيغ فسوف نصل إلى نسب t خاطئة مما يجعل اختباراتنا للفرضيات حول قيم المعلمات في نموذجنا خاطئة أيضا. ونتيجة لذلك فقد نقبل - مثلا - قيمة مقدرة لإحدى المعلمات على أساس أنها معنوية إحصائيا وهي، في الحقيقة، ليست مختلفة معنويا عن الصفر.

نموذج للانحدار الذاتي

لجعل مناقشتنا أكثر علمية دعنا نفترض أن عملية الانحدار الذاتي (أي الطريقة التي يرتبط بها خطأ عشوائي بآخر) تأخذ الشكل التالي:

$$u_t = \gamma u_{t-1} + \varepsilon_t, \quad t = \dots, -2, -1, 0, 1, 2, \dots, \quad (6.18)$$

حيث إن ε_t هو متغير عشوائي موزع توزيعا طبيعيا وله قيمة متوسطة صفرية $E(\varepsilon_t) = 0$ ومستقل عن قيمته في أي فترة زمنية أخرى، لذا يكون $\text{cov}(\varepsilon_t, \varepsilon_s) = 0$ وله تباين ثابت $E(\varepsilon_t^2) = \sigma_\varepsilon^2$. ونفترض أيضا، (لأسباب سنعرضها فيما بعد) أن قيمة γ المطلقة أقل من الواحد الصحيح $|\gamma| < 1$. وباختصار، نفترض أن قيمة الخطأ العشوائي في أي فترة زمنية مرتبطة بالقيمة السابقة له مباشرة بوساطة نموذج انحدار خطي بسيط. لاحظ أنه يفترض أن المعادلة (6.18) تظل قائمة بالنسبة لجميع الفترات الزمنية الماضية والمستقبلية. وفي هذا النموذج، ترتبط قيم الخطأ العشوائي المتتابعة (أي u 's) ببعضها البعض، وإن كان الارتباط ليس تاما وبالتحديد، إذا كانت γ موجبة فإن قيمة u_t سوف تكون مرتبطة إيجابيا مع قيمتها السابقة مباشرة، أما u_{t-1} أما

إذا كانت γ سالبة فإن هذا الارتباط يكون سالبا، وتناظر الحالة الأخيرة مثالنا السابق عن الفرد الذي ينفق أكثر من المعتاد في إحدى الفترات، الفترة t مثلاً، فسيحاول تعويض ذلك بانفاق أقل من المعتاد في الفترة اللاحقة. لاحظ من المعادلة (6.18) أن الفرد، مع ذلك، قد لا يقلل من الانفاق في الفترة التالية $(t+1)$ لحدث آخر غير متوقع قد يجعله يزيد مرة أخرى انفاقه عن القدر المعتاد (أي ε_{t+1}).

سنبين أولاً أنه، في ظل نموذج الانحدار الذاتي في المعادلة (6.18)، فإن $E(u_t) = 0$ لاحظ من المعادلة (6.18) أن:

$$u_{t-1} = \gamma u_{t-2} + \varepsilon_{t-1}, \quad (6.19)$$

وبإحلال قيمة المتغير u_{t-1} من المعادلة (6.19) في المعادلة (6.18) نحصل على:

$$u_t = \gamma(\gamma u_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \varepsilon_t + \gamma \varepsilon_{t-1} + \gamma^2 u_{t-2}. \quad (6.20)$$

ولما كانت المعادلة (6.18) تظل قائمة لجميع الفترات الزمنية، فإحلال ما يعادل u_{t-2} و u_{t-3} وهلم جرا، فسنحصل على المقدار:

$$u_t = \varepsilon_t + \gamma \varepsilon_{t-1} + \gamma^2 \varepsilon_{t-2} + \gamma^3 \varepsilon_{t-3} + \dots, \quad (6.21)$$

الذي لا يتضمن قيماً مبطأة لـ u لأن معامل γ سيكون مرفوعاً لقوى لانهاية (تساوي الصفر). وبأخذ القيمة المتوقعة لـ u_t في المعادلة (6.21)، سيكون لدينا:

$$\begin{aligned} E(u_t) &= E(\varepsilon_t + \gamma \varepsilon_{t-1} + \gamma^2 \varepsilon_{t-2} + \gamma^3 \varepsilon_{t-3} + \dots) \\ &= E(\varepsilon_t) + \gamma E(\varepsilon_{t-1}) + \gamma^2 E(\varepsilon_{t-2}) + \gamma^3 E(\varepsilon_{t-3}) + \dots = 0. \end{aligned} \quad (6.22)$$

ولما كانت $E(\varepsilon_s) = 0$ حيث $s = t, t-1, t-2$ فإن القيمة المتوقعة للخطأ العشوائي لاتزال تساوي الصفر. ويمكننا أن نرى، أنه بافتراض، كما يحدث عادة، استقلالية ε_t عن قيمة المتغير المستقل، X_t مثلاً، في جميع الفترات الزمنية (أي تكون ε_t مستقلة عن X_t لكل s, t)، فإن u_t ستكون مستقلة عن X_t ، وهي هنا غير مرتبطة بها. أي أنه، باستخدام المعادلة (6.21)، ينبغي أن تكون قادراً على إثبات أن:

$$\text{cov}(u_t, X_t) = E(u_t, X_t) = 0$$

وعلى سبيل إشارة يمكن الرجوع إليها مستقبلاً، نلاحظ أنه، طالما أن u_t توليفة خطية من $\varepsilon_t, \varepsilon_{t-1}, \dots$ وطالما أن جميع ε 's مستقلة عن بعضها بعضاً، فإن تباين u_t هو: *

$$\begin{aligned}\sigma_u^2 &= \sigma_\varepsilon^2 + \gamma^2 \sigma_\varepsilon^2 + (\gamma^2)^2 \sigma_\varepsilon^2 + (\gamma^3)^2 \sigma_\varepsilon^2 + \dots \\ &= \sigma_\varepsilon^2 [1 + \gamma^2 + (\gamma^2)^2 + (\gamma^2)^3 + \dots] \\ &= \frac{\sigma_\varepsilon^2}{1 - \gamma^2} = \sigma_u^2,\end{aligned}\quad (6.23)$$

طالما أن جميع ε 's لها التباين نفسه، وبافتراض أن $|\gamma| < 1$. من المعادلة (6.23)، نجد أن تباين u_t لا يتضمن t ، وكما هو الحال لـ ε 's، فإن جميع u_t 's يكون لها التباين نفسه $\sigma_u^2 = \sigma_{u_s}^2 = \sigma_u^2$. ويمكننا أن نرى، أيضاً، من المعادلة (6.23) لماذا نفترض أن $|\gamma| < 1$. فبدون هذا الافتراض، لانتقارب السلسلة في المعادلة (6.23) ويصبح تباين u_t مالاً لانهاية.

دعنا نفحص بعد ذلك التغيرات للأخطاء العشوائية، فبالتعويض عن قيمة u_t من المعادلة (6.18) واستخدام المعادلة (6.23)، نحصل على:

$$\begin{aligned}E(u_t u_{t-1}) &= E[(\gamma u_{t-1} + \varepsilon_t) u_{t-1}] \\ &= \gamma E(u_{t-1}^2) + E(\varepsilon_t u_{t-1}) \\ &= \gamma E(u_{t-1}^2) + 0 = \gamma \sigma_u^2,\end{aligned}\quad (6.24)$$

لأن u_{t-1} باستخدام المعادلة (6.21) يعتمد، فقط، على ε_{t-1} وقيمته المبثأة الأخرى ويعني هذا أن u_{t-1} و ε_t مستقلتان عن بعضهما بعضاً، ولذا فإن:

$$E(\varepsilon_t u_{t-1}) = \text{COV}(\varepsilon_t, u_{t-1}) = 0.$$

* استخدمنا هنا النظرية الأساسية لجمع المتوالية الهندسية التي تبين أنه إذا كان:

$$s = a(1 + \alpha + \alpha^2 + \alpha^3 + \dots), \quad |\alpha| < 1$$

$$s = \frac{a}{1 - \alpha}$$

فإن

فإذا كانت $\alpha \geq 1$ فإن المتوالية لن تتقارب.

وهكذا، نجد من المعادلة (6.24) أن شكل الانحدار الذاتي في معادلة (6.18) يخالف افتراضنا بأن التغير بين الأخطاء العشوائية يساوي الصفر طالما أن γ لا تساوي الصفر.

تأثيره على تباينات المقدرات

سنبين الآن أن الارتباط الذاتي يؤدي إلى صيغ مختلفة لتباين المقدرات. افترض أن لدينا نموذجاً من متغيرين من الشكل الموضح في الفصل الثاني:

$$Y_t = b_0 + b_1 X_t + u_t. \quad (6.25)$$

تذكر أنه وفقاً لافتراضات النموذج (بما فيها التغير الصفري بين الأخطاء العشوائية) وجدنا أن التباين الشرطي للمقدر \hat{b}_1 هو:

$$\text{var}(\hat{b}_1) = \frac{\sigma_u^2}{\sum (X_t - \bar{X})^2}. \quad (6.26)$$

وقد استخدمنا، لاشتقاق هذه النتيجة، المعادلة (2.71) التي تشير إلى أن \hat{b}_1 يمكن التعبير عنه كمايلي:

$$\hat{b}_1 = b_1 + \frac{\sum (X_t - \bar{X}) u_t}{\sum (X_t - \bar{X})^2}. \quad (6.27)$$

عندئذ، دع $A = \sum (X_t - \bar{X})^2$ و $w_t = (X_t - \bar{X})$ ، ومن ثم، نعيد كتابة (6.27) في شكل مفصل على النحو:

$$\hat{b}_1 = b_1 + \frac{w_1}{A} u_1 + \dots + \frac{w_n}{A} u_n. \quad (6.28)$$

وعلى افتراض أن قيم X 's معطاة (ومن ثم، w 's و A) فإن \hat{b}_1 تكون توليفة خطية من الأخطاء العشوائية. عند ذلك نشق المعادلة (6.26) من المعادلة (6.28) باستخدام صيغة التباين لمجموع خطي من المتغيرات العشوائية غير المترابطة. ولكن، إذا كانت هناك مشكلة ارتباط ذاتي فليس بإمكاننا استخدام هذه الصيغة لأن u_t وبالتالي، الحدود الموجودة في المعادلة (6.28) مترابطة. ويعني هذا أن المعادلة (6.26) لم تعد

الصيغة الصحيحة لتباين \hat{b}_1 . ونتيجة لذلك فإن استخدام صيغنا التقليدية لاختبار الفرضيات لم تعد صحيحة كذلك.*

الوسط الحسابي للمقدرات

من السهل أن نرى أن الارتباط الذاتي لا يؤدي إلى تحيز في \hat{b}_1 حيث ما يزال الخطأ العشوائي u_i ، في ظل افتراضاتنا السابقة، مستقلاً عن جميع قيم X_i ، ومن ثم، غير مرتبط بها (ومن ثم، w_i) وماتزال قيمته المتوقعة هي الصفر. فإذا أخذنا القيمة المتوقعة لـ (6.28) لأي قيم معطاة لـ X_i فإننا نحصل على:

$$E(\hat{b}_1) = b_1 + \frac{w_1}{A} E(u_1) + \dots + \frac{w_n}{A} E(u_n) = b_1. \quad (6.29)$$

وبالمثل، فإنه، باستخدام القاعدة $\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}$ ينبغي أن تكون قادراً على إثبات:

$$\begin{aligned} E(\hat{b}_0) &= b_0 + b_1 \bar{X} + E(\bar{u}) - \bar{X} E(\hat{b}_1) \\ &= b_0 \end{aligned} \quad (6.30)$$

هنا ينبغي أن تكون السمة العامة لمشكلة الارتباط المتسلسل الذاتي واضحة، فعندما توجد هذه المشكلة فسيكون لدينا تغير منتظم في قيم الخطأ العشوائي للملاحظات المتتالية. مثل هذا النمط من التغير لا يؤدي إلى إيجاد مقدرات متحيزة للمعلمات. إلا أن صيغنا للتباين لم تعد صحيحة، ونتيجة لذلك، وبدون نتائج إضافية، فإنه لا يمكننا اختبار الفرضيات وإنشاء فترات ثقة. ومن الواضح أننا نحتاج إلى خصائص مرغوب فيها في منهجنا للتقدير. وأكثر من ذلك، فمن البديهي، طالما أن منهجنا في التقدير لا يأخذ في الحسبان صراحة الارتباط المتسلسل الذاتي فإنه قد لا يعطينا أكثر المقدرات دقة للمعلمات. أي أنه، إذا كان هناك نمط محدد للتغير بين الأخطاء

* باستخدام المعادلة (6.28)، ينبغي أن تكون لديك القدرة على إثبات أنه، في حالة الارتباط الذاتي، فإن تباين

b (أي $E(\hat{b}_1 - b_1)^2$) يتضمن القيمة المتوقعة لحدود ضرب التقاطع في المعادلة (6.28). وفي حالة غياب الارتباط

الذاتي، تكون جميع حدود التقاطع هذه صفرية وبذلك، تسقط، تاركة لنا الصيغة (6.26) لتباين \hat{b}_1 .

العشوائية، فينبغي أن نكون قادرين على الحصول على تقدير وتنبؤ أفضل بأخذ هذه المعلومات الإضافية في حساباتنا. وسوف نفحص الآن كيف يمكن أن يحدث ذلك.

طريقة تقدير معممة

افترض أن نموذجنا يتكون من:

$$Y_t = b_0 + b_1 X_t + u_t \quad (6.31)$$

$$u_t = \gamma u_{t-1} + \varepsilon_t, \quad (6.32)$$

حيث $|\gamma| < 1$ و ε_t تحقق الافتراضات المعتادة السابق تكوينها كافة. مشكلتنا هنا تنحصر في كيفية استخدام المعلومات المعطاة في المعادلة (6.32) لتحسين مقدراتنا لمعاملات المعادلة (6.31).

افترض مبدئياً أن قيمة γ معلومة. فإذا أخذنا الصيغة المبطأة بالمعادلة (6.31) وضربناها بوساطة γ نحصل على:

$$\gamma Y_{t-1} = \gamma b_0 + \gamma b_1 X_{t-1} + \gamma u_{t-1}. \quad (6.33)$$

وبطرح المعادلة (6.33) من المعادلة (6.31) نحصل على:

$$Y_t - \gamma Y_{t-1} = (b_0 - \gamma b_0) + (b_1 X_t - \gamma b_1 X_{t-1}) + (u_t - \gamma u_{t-1}). \quad (6.34)$$

وبإعادة ترتيب الحدود في المعادلة (6.32)، نجد أن:

$$\varepsilon_t = (u_t - \gamma u_{t-1}),$$

التي، عند التعويض عن الحد الأخير في المعادلة (6.34)، نحصل على:

$$Y_t - \gamma Y_{t-1} = (b_0 - \gamma b_0) + (b_1 X_t - \gamma b_1 X_{t-1}) + \varepsilon_t. \quad (6.35)$$

ويمكننا إعادة كتابة المعادلة (6.35) نتحصل على:

$$\begin{aligned} Y'_t &= B + b_1 X'_t + \varepsilon_t, \\ Y'_t &= Y_t - \gamma Y_{t-1}, \\ B &= b_0 - \gamma b_0, \\ X'_t &= X_t - \gamma X_{t-1}. \end{aligned} \quad (6.36)$$

لاحظ أننا، عند ترجمة المعادلة (6.31) إلى (6.36)، خسرنا مشاهدة واحدة بسبب الإبطاء والطرح في المعادلة (6.34).

وهكذا، تصبح المعادلة (6.31) الشكل المعتاد لنموذج الانحدار، خصوصا أن ε_t (وليس u_t) يحقق الافتراضات المرتبطة بخصائص الخطأ العشوائي كافة. لذلك، يمكننا أن نتبع منهج التقدير الذي وضعناه من قبل: نضع الشروط

$$\sum_{t=2}^n (X'_t \hat{\varepsilon}_t) = 0 \text{ و } \sum_{t=2}^n \hat{\varepsilon}_t = 0$$

وبتلك الوسيلة، يمكننا اشتقاق معادلتين طبيعيتين،

وبحلها، نحصل على مقدراتنا لـ $\hat{\beta}$ و \hat{b}_1 ، للمعلمتين في المعادلة (6.36). ويمكننا بعدئذ أن نقدر b_0 بوساطة:

$$\hat{b}_0 = \frac{\hat{B}}{1-\gamma}. \quad (6.37)$$

ويمكن، أيضا، إثبات أن:

$$\text{var}(\hat{b}_0) = \left(\frac{1}{1-\gamma} \right)^2 \text{var}(\hat{B}), \quad (6.38)$$

طالما أن \hat{b}_0 مرتبطة ارتباطا خطيا تاما مع $\hat{\beta}$. وطالما أن ε_t يحقق افتراضاتنا المعتادة كافة، فإنه يمكن الوصول إلى تباينات $\hat{\beta}$ و \hat{b}_1 بوساطة صيغنا المعتادة:

$$\text{var}(\hat{B}) = \frac{\sigma_t^2 \sum_{t=2}^n (X'_t)^2}{n' \sum_{t=2}^n (X'_t - \bar{X}')^2}, \quad \text{var}(\hat{b}_1) = \frac{\sigma_t^2}{\sum_{t=2}^n (X'_t - \bar{X}')^2} \quad (6.38)$$

حيث إن $n' = n-1$ ، نظرا لفقدان إحدى المشاهدات في عملية الإبطاء والطرح للحصول على X'_t .

يمكننا الآن - من حيث المبدأ، في الأقل - اختبار الفرضيات وإنشاء فترات ثقة صحيحة إذا ما قدرنا معلمتنا الأساسية بوساطة المعادلة (6.36). ونشير، أيضا،

إلى أن المقدرات التي حصل عليها من المعادلة (6.36) كفاء، ويعني ذلك بدهيا*، أنه إذا كان حجم العينة كبيرا فإن تباينات مقدرات (6.36) ستكون أقل من تباينات أي مقدرات أخرى غير متحيزة لـ B و b_1 أو تساويها. والسبب الذي يمنعنا من القول أن مقدراتنا سوف يكون لها أقل التباينات بغض النظر عن حجم العينة هو أن طريقتنا في التقدير تتضمن، فقط، أن واحدة من المشاهدات. وأساساً، تتضاءل أهمية هذه المشاهدة بتزايد حجم العينة.**

وهكذا، فقد وجدنا طريقة للتقدير بصفات مرغوب فيها لدمج المعلومات المتاحة حول العلاقة بين الأخطاء العشوائية ذاتها في منهجنا للتقدير. وبالتحديد، فقد حولنا نموذج الانحدار الذي يعاني من الارتباط الذاتي إلى نموذج يحقق كافة الافتراضات المرتبطة بنموذج الانحدار الأساسي (ومن بينها التغير الصفري بين الأخطاء العشوائية) وبعد ذلك نطبق طرقنا المعتادة في التقدير. والصعوبة التي نواجهها في تطبيق هذه الطريقة هي أن γ (وعموماً) غير معلومة، ولذا ينبغي علينا أولاً أن نقدر قيمة γ .***

ويمكننا أداء ذلك من خلال رؤية البواقي من معادلتنا الأصلية (6.31). ويتذكر أن $E(u_t) = 0$ وأن $\text{cov}(u_t, X_t) = 0$ يمكننا اشتقاق مقدرات غير متحيزة للمعلمات في

* ليس هذا تعريفاً اصطلاحياً للمقدر الكفاء، فمثل هذا التعريف خارج عن نطاق هذا الكتاب وغير ضروري لفهم النتائج التي سترد فيها بعد.

** تعرف الكفاءة، عادة، بدلالة ما يسمى بمتوسط مربع الخطأ mean square error (م م خ). أي افترض

أن $\hat{\alpha}$ هي مقدر لـ α . حيث، فإن م م خ لـ $\hat{\alpha}$ يساوي مجموع تباينها $\sigma_{\hat{\alpha}}^2$ ومربع تحيزها $E(\hat{\alpha} - \alpha)^2$.

وفي حالتنا المذكورة أعلاه، طالما أن مقدراتنا غير متحيزة فإن م م خ يساوي التباين فقط. وعلى أي الأحوال، ويتجاهل قليل من الأشياء غير المهمة، إذا كانت $\hat{\alpha}$ مقدرًا كفئًا لـ α فإننا نتوقع (في حالة العينات الكبيرة) أن م م خ لـ سوف يكون أقل من م م خ لأي مقدر متنسق لـ α أو يساويه.

*** نلاحظ في معادلات الانحدار أنه جميع متغيراتنا تأخذ شكل الفروق من الدرجة الأولى First difference

form، أي بالنسبة لحالة المتغيرين الاثنين فقط $Y_t - Y_{t-1}$ يتم انحدارها على $X_t - X_{t-1}$. ومن المعادلة

(6.36)، يمكن أن يتبين لنا أن مثل هذه المعادلة هي حالة خاصة لنموذج الانحدار الذاتي حيث تكون γ

مساوية الواحد. لاحظ من المعادلة (6.23)، أنه بالنسبة لهذا النموذج، فإن تباين u سيكون لانهائياً. وأكثر

من ذلك يكون هذا المنهج مقيداً جداً، طالما أن نتائجه مبنية على تحقق الشرط $\gamma=1$.

المعادلة (6.31) عن طريق فرض الشروط $\sum_{i=1}^n \hat{u}_i = 0$ و $\sum_{i=1}^n (\hat{u}_i X_i) = 0$ ويعطينا هذه المعادلات الطبيعية المعتادة.

$$\begin{aligned} \sum Y_i &= n\hat{b}_0 + \hat{b}_1 \sum X_i, \\ \sum (X_i Y_i) &= \hat{b}_0 \sum X_i + \hat{b}_1 \sum X_i^2, \end{aligned} \quad (6.40)$$

والتي يمكننا حلها للحصول على المقدرات غير المتحيزة \hat{b}_0 و \hat{b}_1 . ويمكننا، حينئذ، أن نستخدم \hat{b}_0 و \hat{b}_1 للحصول على مقدر (\hat{u}_i) لقيمة الخطأ العشوائي:

$$\hat{u}_i = Y_i - \hat{Y}_i = Y_i - (\hat{b}_0 + \hat{b}_1 X_i). \quad (6.41)$$

ولتقدير γ نعوض ببساطة عن قيمة (\hat{u}_i) من المعادلة (6.41) في العلاقة المقترحة في المعادلة (6.32)، أي:

$$\hat{u}_i = \gamma \hat{u}_{i-1} + \varepsilon_i. \quad (6.42)$$

اعتبر المعادلة (6.42) نموذجاً للانحدار. وبما أن ε_i مستقلة عن u_i دعنا نجعل ε_i غير مرتبطة بـ u_{i-1} .^{*} وبلاستعانة بهذا الفرض يمكننا أن نقدر γ في المعادلة (6.42) بوساطة منهجنا العادي. وبالتحديد (على سبيل المراجعة) يمكننا إعادة كتابة المعادلة (6.42) على النحو التالي:

$$\hat{u}_i = \gamma \hat{u}_{i-1} + \hat{\varepsilon}_i, \quad (6.43)$$

حيث إن $\hat{\varepsilon}_i = (\hat{u}_i - \gamma \hat{u}_{i-1})$ هو مقدر الخطأ العشوائي و $\hat{\gamma}$ هو مقدرنا لـ γ . ويعني افتراضنا $\text{cov}(\varepsilon_i, \hat{u}_{i-1}) = 0$ أننا نفرض الشرط التالي (تذكر أن إحدى المشاهدات تفقد

* يتضائل الترابط بين ε_i و u_{i-1} ويؤول إلى الصفر، كلما ازداد حجم العينة إلى ما لا نهاية. أي أنه كلما كان حجم العينة صغيراً كان هنالك ترابط بين ε_i و \hat{u}_{i-1} وذلك بسبب أن \hat{u}_{i-1} تعتمد على كل من \hat{b}_0 و \hat{b}_1 اللذين يعتمدان، بدورهما، على جميع ε_i مشتملة على ε_i . ولكن، مع زيادة حجم العينة، وطالما أن \hat{b}_0 و \hat{b}_1 مستقرتان، فإنهما سوف تتقاربان في الاحتمال إلى b_0 و b_1 لذا، فإن \hat{u}_{i-1} سوف تنحرف في النهاية عن u_{i-1} باحتمال يساوي الصفر. ويمكن افتراض صحة المعادلة (6.42)، باحتمال قدره الواحد الصحيح، فقط، في حالة ما إذا كان حجم العينة لانهائياً. باختصار، ينبغي علينا أن نعد (6.42) معادلة تقريبية (أوللعينات الكبيرة).

وللقيام بذلك نضرب حدود المعادلة (6.43) بوساطة \hat{u}_{t-1} ، ثم

نجمع على مدى العينة. وأخيرا نطبق شرطنا للحصول على:

$$\sum_{t=2}^n (\hat{u}_t u_{t-1}) = \hat{\gamma} \sum_{t=2}^n (\hat{u}_{t-1})^2 + \sum_{t=2}^n (\hat{\varepsilon}_t \hat{u}_{t-1}) = \hat{\gamma} \sum_{t=2}^n (\hat{u}_{t-1})^2. \quad (6.44)$$

ومن المعادلة (6.44) يكون مقدرنا لـ γ هو:

$$\hat{\gamma} = \frac{\sum_{t=2}^n (\hat{u}_t \hat{u}_{t-1})}{\sum_{t=2}^n (\hat{u}_{t-1})^2} \quad (6.45)$$

يمكننا الآن أن نستخدم المنهج الذي وضعناه من قبل (باستخدام $\hat{\gamma}$ بدلا من γ) للحصول على مقدرات لنموذجنا للانحدار. في حالة نموذج من متغيرين، فسيكون لدينا:

$$\hat{b}_1 = \frac{\sum_{t=2}^n (X_t^* - \bar{X}) Y_t^*}{\sum_{t=2}^n (X_t^* - \bar{X}^*)^2} \quad (6.46)$$

وأیضا

$$\hat{b}_0 = \frac{\hat{B}}{1 - \hat{\gamma}},$$

$$X_t^* = X_t - \hat{\gamma} X_{t-1}, \quad (6.47)$$

$$Y_t^* = Y_t - \hat{\gamma} Y_{t-1}, \quad \text{وحيث}$$

$$\hat{B} = \bar{Y}^* - \hat{b}_1 \bar{X}^*.$$

وبالمثل ووفقا لمناقشتنا اللاحقة ستكون صيغتنا التباين هي:

$$\begin{aligned} \text{var}(\hat{b}_0) &= \frac{1}{(1-\hat{\gamma})^2} \text{var}(\hat{B}), \\ \text{var}(\hat{b}_1) &= \frac{\sigma_\varepsilon^2}{\sum_{i=2}^n (X_i^* - \bar{X}^*)^2} \end{aligned} \quad (6.48)$$

حيث

$$\text{var}(\hat{B}) = \frac{\sigma_\varepsilon^2 \sum_{i=2}^n (X_i^*)^2}{n' \sum_{i=2}^n (X_i^* - \bar{X}^*)^2} \quad (6.49)$$

حيث $n' = n-1$. وباختصار طالما أننا حولنا معادلتنا باستخدام $\hat{\gamma}$ ، فإننا سوف نعامل $\hat{\gamma}$ كما لو أنها γ وسنستخدم جميع صيغنا المعتادة. وهذا يتضمن تقدير σ_ε^2 في صيغ التباين المذكورة أعلاه. وبالتحديد ينبغي أن يكون واضحاً من (6.36) أن مقدرنا σ_ε^2 سيكون:

$$\hat{\sigma}_\varepsilon^2 = \sum_{i=2}^N \frac{(Y_i^* - \hat{B} - \hat{b}_1 X_i^*)^2}{(n' - 2)} \quad (6.49)$$

لنبحث الآن قضايا اختبار الفرضيات وتكوين فترات الثقة. من (6.45) نجد أن $\hat{\gamma}$ تعتمد على الأخطاء العشوائية المقدرة. ومن (6.46) نجد أن \hat{b}_1 تعتمد اعتماداً غير خطي على $\hat{\gamma}$ وينتج عن ذلك أن \hat{b}_1 تعتمد - من ضمان أشياء أخرى - على الأخطاء العشوائية المقدرة، إضافة إلى ذلك وطالما أن الأخطاء العشوائية [انظر المعادلة و (6.25)]، فإن \hat{b}_1 تعتمد اعتماداً غير خطي على الأخطاء العشوائية ونتيجة لذلك، فإن \hat{b}_1 ليست موزعة توزيعاً طبيعياً، كما أن النسبة $(\hat{b}_1 - b_1) / \hat{\sigma}_{b_1}$ ليست متغير t بدرجات حرية $n-2$ ، حيث إن $\hat{\sigma}_{b_1}$ هو الانحراف المعياري المقدّر لـ \hat{b}_1 . ويمكن الوصول إلى نتائج مشابهة لكل من \hat{B} و \hat{b}_0 طالما أنهما يعتمدان اعتماداً غير خطي أيضاً، على $\hat{\gamma}$.

ولحسن الحظ، يمكن اختبار الفرضيات بطريقة تقريبية أو تكون فترات ثقة تقريبية عن طريق افتراض أن النسب $(\hat{b}_0 - b_0)/\hat{\sigma}_{\hat{b}_0}$ و $(\hat{b}_1 - b_1)/\hat{\sigma}_{\hat{b}_1}$ وأخير $(\hat{B} - B)/\hat{\sigma}_B$ هي متغيرات طبيعية معيارية. فإذا كان حجم العينة لانهائياً، فإن هذه النسب، في الحقيقة هي متغيرات طبيعية معيارية. وفي حالة العينة النهائية (المحدودة) يمكن أن ينظر إلى افتراض الطبيعية على أنه افتراض تقريبي. ولذلك، فإن اختبار الفرضيات وتكوين فترات الثقة المبنية على افتراض الطبيعية ينبغي أن ينظر إليها على أنها تقريبية. ونلاحظ (بالمثل) أن صيغ التباين السابقة (عند استخدام $\hat{\lambda}$ بدلا من λ) هي تقريبية، أيضا، بمعنى أنها تكون صحيحة فقط في حال، العينة اللانهائية.

يتبع عن المناقشة السابقة أن فترة ثقة تقريبية 95% لـ b_1 ستكون $(\hat{b}_1 \pm 1.96 \sigma_{\hat{b}_1})$ فإذا كنا نرغب في اختبار الفرضية $H_0 : b_1 = 0$ مقابل الفرضية البديلة $H_1 : b_1 \neq 0$ عند مستوى معنوية 5%. فإننا سنقبل فرضية العدم H_0^* إذا كان $|\hat{b}_1 / \hat{\sigma}_{\hat{b}_1}| < 1.96$ ونرفضها إذا لم تتحقق هذه المتباينة. أما إذا كان الاهتمام بما نستمر في تسميته نسبة ونفرضها إذا لم تتحقق هذه المتباينة. فإن الاختلاف الوحيد هو أن القيمة الحرجة الدقيقة ترتبط بالتوزيع الطبيعي بدلا من توزيع t . ولهذا السبب نجد الباحثين يحسبون ويكتبون نتائج الانحدار غالبا في شكل نسب تسهيلات لقراءتهم.

ونشير هنا إلى أن مقدراتنا المبنية على $\hat{\gamma}$ لم تعد غير متحيزة، ولكنها تتسم بصفة مرغوب فيها وهي الاتساق^{**}، إضافة إلى ذلك، فإن هذه المقدرات تتسم بالكفاءة، لذلك لا توجد مقدرات متسقة أخرى أفضل لمعلومات النموذج (في الأقل في العينات الكبيرة).

* لا نستطيع (كما يذكر المؤلف) القول بقبول فرضية العدم (اصطلاحيا وإنما القبول، فقط، يكون للفرضية البديلة، ويمكننا، في حالة عدم تحقق الفرضية البديلة، أن نذكر أننا لم نستطيع رفض فرضية العدم، وهناك فرق واضح بين عدم القدرة على رفض فرضية العدم وبين قبولها. ملحوظة المترجم.

** لمعرفة المزيد من القضايا المتضمنة في التحليل السابق يرجى الرجوع إلى:

Arthur S. Goldberger, *Econometric Theory* (New York: Wiley, 1964) Chap. 6.

حال نموذج الانحدار المتعدد

يمكن، بسهولة، توسيع نطاق الطريقة السابقة ليشتمل على معالجة مشكلة الارتباط الذاتي في حالة الانحدار المتعدد. افترض (مثلاً) أننا قد احتفظنا بالافتراضات السابقة كافة ماعدا أن لدينا الآن متغيرات مستقلة عددها k ، حيث نأخذ نموذجنا الشكل:

$$Y_t = b_0 + b_1 X_{1t} + \dots + b_k X_{kt} + u_t, \quad (6.50)$$

$$u_t = \gamma u_{t-1} + \varepsilon_t.$$

سنقدر أولاً معلمات النموذج (b_0, b_1, \dots, b_k) بطريقتنا المعتادة ثم نقدر بعد ذلك الخطأ العشوائي بوساطة $\hat{u}_t = Y_t - \hat{Y}_t$. ثم نقدر γ بعد ذلك بوساطة (6.45)، ونحول بعدها المتغير التابع إلى $Y_t^* = Y_t - \gamma \hat{Y}_{t-1}$ ، وكذلك المتغيرات المستقلة إلى $X_{it}^* = X_{it} - \gamma \hat{X}_{i(t-1)}$ لدينا نموذج الانحدار التالي:

$$Y_t^* = B + b_1 X_{1t}^* + \dots + b_k X_{kt}^* + \varepsilon_t, \quad (6.51)$$

وسوف نقدر B, b_1, \dots, b_k وأيضاً، تباينات مقدراتنا بطرق التقدير المعروفة، ومرة أخرى، فإن مقدرات المعلمات ستكون متحيزة إلا أنها متسقة. وأخيراً سنختبر الفرضيات أو فترات الثقة عن طريق افتراض أن النسب $(\hat{b}_i - b_i) / \hat{\sigma}_{\hat{b}_i}$ هي متغيرات طبيعية معيارية، وكما عرفنا في الحالة السابقة، فإن اختبارات الفرضيات هذه وتكوين فترات الثقة أو حتى صيغ التباين لمقدرات المعلمات تكون صحيحة تماماً في حالة العينة اللانهائية. ولذا، ينبغي أن نفسر نتائجنا بأنها تقريب للعينات النهائية. ونشير في هذه العجالة إلى أن هذه الطريقة التي شرحناها ليست هي الطريقة الوحيدة لتصحيح الارتباط الذاتي. فهناك طريقتان تستخدمان استخداماً واسعاً وهما طريقتا كوكرين أوركوت Cochrane-Orcutt وهيلدروث لو Hildreth-Lu. وبالنسبة للنموذج الحالي، فهذه الطرق تعادل الطريقة التي اتبعناها من حيث إن هذه الطرق تنتج مقدرات تكون خواص العينات الكبيرة منها متماثلة مع مقدراتنا، أي أن المقدرات تتسم بالاتساق والكفاءة.

اختبار ديربن - واتسون للارتباط الذاتي

دعنا نفترض أنه إذا كانت الأخطاء العشوائية في نموذج الانحدار مرتبطة ذاتياً، فإن علاقة الارتباط هذه تأخذ الشكل الموجود في النموذج (6.32). ولدينا الآن طريقة لمعالجة مشكلة الارتباط الذاتي. ولكننا لم نوجد الوسائل التي نتعرف بها على ما إذا كان لدينا مشكلة ارتباط ذاتي أم لا. بدلا من ذلك كان منهجنا السابق يعتمد على افتراض أنه في (6.32) تكون $\gamma \neq 0$ ، ومن الواضح أن من الأفضل أن نختبر هذه الفرضية.

وإحدى الطرق المباشرة لأداء ذلك هو أخذ $\gamma = 0$ على أنها فرضية العدم ثم فحص إمكانية رفض هذه الفرضية في صالح الفرضية البديلة $\gamma \neq 0$ عند مستوى من المعنوية. ونشئ فترة للثقة (بما يماثل طريقتنا في الفصل الثالث) لمقدرنا لـ γ . فإذا تضمنت فترتنا للثقة لـ $\hat{\gamma}$ الصفر، فسوف نقبل فرضية العدم بأن $\gamma = 0$ ، وإذا لم تتضمنه (بسبب أن $\hat{\gamma}$ موجبة أو سالبة بدرجة كافية) فسنرفض فرضية أن $\gamma = 0$ ، وفي هذه الحالة الأخيرة، نكون قد قبلنا فرضية أن الخطأ العشوائي مرتبط ذاتياً.* يوجد لحسن الحظ اختبار لهذا النوع من الارتباط الذاتي طور ج. ديربن - و ج. واتسون**، وبستخدم الاختبار ما يشار إليه عادة باحصائية d لديربن واتسون، والمبنية على مجموع مربع الفروق في القيم المتتابة للأخطاء العشوائية المقدرة.

$$d = \frac{\sum_{t=2}^n (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^n \hat{u}_t^2} \quad (6.52)$$

* لمناقشة هذه الطرق أنظر:

S. Goldfeld and R. Quandt. *Non-Linear Methods in Econometrics* (Amsterdam: North Holland, 1972), pp. 183-186.

** يرجع إلى:

J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least-Squares Regression", parts I and II, *Biometrika* 37 (1950), pp. 409-428 and 38 (1951), pp. 159-178. See also the discussion of the test in Arthur S. Goldberger, *Econometric Theory* (New York: Wiley, 1964), pp. 243-244; and in J. Johnston. *Econometric Methods*, 2nd ed. (New York: McGraw-Hill, 1972), pp. 249-254.

وبديهيًا يمكننا أن نرى أنه إذا كان لدينا ارتباط ذاتي موجب فإن القيم المتتابعة للأخطاء العشوائية سوف تميل للاقتراب من بعضها بعضًا اقترابًا غير عادي، فقيمة موجبة للخطأ العشوائي في الفترة t سوف تتبعها، على الأرجح، قيمة موجبة أخرى في الفترة $t+1$. ويعني هذا أن الحدود في بسط (6.52) سوف تكون صغيرة نسبيًا ولذلك نتوقع أن الارتباط الذاتي الموجب ينتج عنه قيم صغيرة لـ d . وعلى العكس فإن الارتباط الذاتي السالب سيؤدي إلى إيجاد اختلافات كبيرة بين القيم المتتابعة لـ u_t . وتكون علامة هذا النوع من الارتباط الذاتي هي قيمة كبيرة غير عادية لـ d .

دعنا نفترض صحة افتراضنا في نموذج الانحدار الأصلي بأن $E(u_s u_t) = 0$ ، حيث $s \neq t$ ، لذا لا توجد مشكلة الارتباط الذاتي. في هذه الحال، نتوقع أن التباين بين البواقي المقدرة \hat{u}_t و \hat{u}_{t-1} يكون صفرًا بالتقريب. عندما يكون ذلك صحيحًا فإنه يتبين لنا عن طريق فك بسط احصائية d في (6.52) أنه إذا كانت n كبيرة، فإن d ينبغي أن تكون قريبة من 2.*

$$d = \frac{\sum_{t=2}^n (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^n \hat{u}_t^2} = \frac{2 \sum_{t=2}^n \hat{u}_t^2 - 2 \sum_{t=2}^n (\hat{u}_t \hat{u}_{t-1})}{\sum_{t=1}^n \hat{u}_t^2} \quad (6.53)$$

$$= \frac{\left[2 \sum_{t=2}^n \hat{u}_t^2 / (n-1) \right] - \left[2 \sum_{t=2}^n \hat{u}_t \hat{u}_{t-1} / (n-1) \right]}{\left[\sum_{t=1}^n \hat{u}_t^2 / (n-1) \right]}$$

طالما أن $\sum_{t=2}^n u_t^2 = \sum_{t=2}^n \hat{u}_t^2$ و $\sum_{t=2}^n u_t^2 = \sum_{t=1}^n \hat{u}_t^2$ ، $\sum_{t=2}^n (\hat{u}_t \hat{u}_{t-1}) / (n-1) = 0$ وحقيقة فإنه وعلى سبيل التعميم إلى حد ما) ينتج من (6.53) أنه إذا كانت n كبيرة. وإذا

* إذا كان حجم العينة n لانهايا، فإن قيمة d ستصبح 2 باحتمال قدره الواحد الصحيح.

افترضنا أن نموذجنا المرتبط ذاتيا السابق، $u_t = \gamma u_{t-1} + \varepsilon_t$

$$d \doteq \frac{2 \text{var}(u_t) - 2 \text{cov}(u_t, u_{t-1})}{\text{var}(u_t)} \quad (6.54)$$

$$= \frac{2\sigma_u^2 - 2\gamma\sigma_u^2}{\sigma_u^2} = 2(1 - \gamma);$$

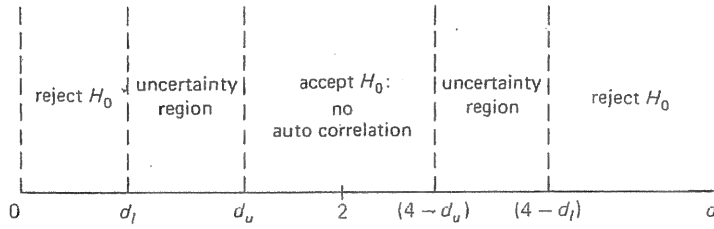
وطالما أن \hat{u}_t مقدر لـ u_t ، ومن (6.42)، $\text{cov}(u_t, u_{t-1}) = \gamma\sigma_u^2$ وباختصار فإننا نرى أن:

$$\begin{aligned} d &\doteq 2 & \gamma &= 0 & \text{توحي بأن} \\ d &\doteq 0, & \gamma &= 1 & \text{توحي بأن} \\ d &\doteq 4, & \gamma &= -1 & \text{توحي بأن} \end{aligned} \quad (6.55)$$

يوحي لنا كل ماسبق أننا إذا أردنا اختبار فرضية العدم وهي عدم وجود ارتباط ذاتي $H_0: \gamma = 0$ إزاء الفرضية البديلة بوجوده $H_1: \gamma \neq 0$ فإننا سنقبل H_0 إذا كانت قيمة d قريبة قريبا كافيا من 2، وسوف نقبل الفرضية البديلة إذا لم تكن d كذلك. وبالمقابل فإن قيم d التي تكون قريبة من الصفر أو من 4 ستقودنا إلى قبول الفرضية البديلة $H_1: \gamma \neq 0$.

ولسوء الحظ، وبسبب وجود خواص إحصائية معينة للإحصائية d ، فإن المشكلة أكثر تعقيدا. وبخاصة أن المناطق المناظرة لقبول أو رفضها فرضية العدم (لإحصائية d) بعدم وجود ارتباط ذاتي لا تستنفذ جميع القيم الممكنة لـ d . لذلك يوجد مدى من القيم لا يمكننا خلالها أن نقبل أو نرفض H_0 . وبالتحديد، في حالة اختبار ديربن واتسون ذو الطرفين مع $H_0: \gamma = 0$ إزاء $H_1: \gamma \neq 0$ ، توجد لدينا مجموعة من خمس مناطق لقيم d كما تظهر في الشكل (٦-٥). فإذا كانت d أقل من d_L أو أكبر من $(4 - d_L)$ فإننا نرفض فرضية العدم في صالح الفرضية البديلة، ويتضمن هذا وجود الارتباط الذاتي. وعلى العكس، إذا كانت قيمة d قريبة من 2، أو بصورة أكثر دقة بين d_u و $(4 - d_u)$ نقبل فرضية العدم بعدم وجود ارتباط ذاتي. أما إذا كانت

قيمة d تنحصر بين d_L و d_u أو بين $(4 - d_u)$ أو $(4 - d_L)$ فإن اختبار ديربن واتسون يصبح غير حاسم لأنه عند هذه القيم من d لا يمكننا عند مستوى محدد من المعنوية، أن نستنتج وجود الارتباط الذاتي أو عدم وجود بين الأخطاء العشوائية. أي أنه، على العكس من اختباراتنا السابقة، يتضمن اختبار ديربن واتسون (وبسبب صعوبات إحصائية معينة) مناطق عدم تأكيد.



الشكل رقم (٥-٦)

وتتبع طريقة اختبار الذيل الواحد مباشرة المنهج السابق. افترض مثلاً أننا مهتمون بالفرضية $H_0: \gamma = 0$ مقابل الفرضية $H_1: \gamma > 0$. حيثد سوف نقبل H_0 إذا كانت d بعيدة «بعداً كافياً» عن الصفر، واصطلاحياً، وبدلالة الشكل (٥-٦) سوف نقبل H_0 إذا كانت $d_u < d$ ونرفضها إذا كانت $d < d_L$ ، ويكون الاختبار غير حاسم إذا كانت $d_L < d < d_u$. وبالمثل إذا كانت الفرضية البديلة هي $\gamma < 0$ ، فإننا سنقبل H_0 إذا كانت $d < 4 - d_u$ ونرفضها إذا كانت $d > 4 - d_L$ ، ويكون الاختبار غير حاسم إذا كانت $4 - d_u < d < 4 - d_L$.

يوجد في نهاية هذا الكتاب جدول بقيم d_L و d_u في الجدول الإحصائي ٤ ولإيجاد قيمة معينة لـ d_L و d_u للمشكلة الحالية، نحتاج لمعرفة مستوى المعنوية، وما إذا كان الاختبار ذا طرف واحد أم اثنين. وكذلك حجم العينة، وأخيراً، عدد

المتغيرات المستقلة (k') في معادلة الانحدار. * وبالرجوع إلى الجدول الإحصائي نجد أنه، على سبيل المثال، وعند مستوى معنوية 5% ($\alpha = 0.05$) في اختبار ذو الطرفين، فإنه إذا كان لدينا 50 مشاهدة ($n=50$)، ومعادلة لها ثلاثة متغيرات مستقلة ($k'=3$)، حينئذ، فإن $d_L = 1.34$ و $d_U = 1.54$. وعند الحصول على هذه الأرقام، نلاحظ أن قيم d_L و d_U المعطاة في الجدول والتي تناظر 0.025 مستوى معنوية (أي 0.025 في كل طرف). في هذه الحال، تكون المناطق الخمس في الشكل (٦-٥) هي:

$$a.(0, d_L) = (0, 1.34),$$

$$b.(d_L, d_U) = (1.34, 1.59),$$

$$c.(d_U, 4 - d_U) = (1.59, 2.41),$$

$$d.(4 - d_U, 4 - d_L) = (2.41, 2.66),$$

$$e..(4 - d_L, 4) = (2.66, 4.00).$$

بعد ذلك، نستخدم المعادلة (6.52) لحساب القيمة الفعلية لـ d من القيم المشاهدة لمتغيراتها ونحدد في أي من المناطق الخمس تقع هذه القيمة لمعرفة ما إذا كان علينا أن نتخوف من وجود مشكلة الارتباط الذاتي.

تطبيق

قد يكون من المفيد لمراجعة معالجتنا للارتباط الذاتي أن ننهي مناقشتنا بمثال توضيحي يشتمل على أرقام واقعية. افترض أننا نرغب في تقدير دالة استهلاك بسيط تأخذ الشكل:

$$C_t = b_0 + b_1 Y_{dt} + u_t, \quad (6.56)$$

حيث إن C_t (كما في المثال السابق) الإنفاق الاستهلاكي و Y_{dt} الدخل المتاح، ولدينا مجموعة من المشاهدات السنوية عن الاستهلاك الاجمالي، والدخل المتاح للولايات

* لا تشمل k على الحد الثابت. ويمكن أن تعرف k' (مثلاً) بأنها عدد معلمات الميل في الانحدار.

المتحدة الأمريكية للسنوات ١٩٥١-١٩٦٩م، التي تظهر في الجدول رقم (١-٦). فإذا أجرينا انحدارا لـ C على Y_d باستخدام البيانات المتاحة في الجدول رقم (١-٦) نحصل على:

$$\hat{C}_t = 3.29 + 0.906Y_{dt}, \quad n = 19, \quad (6.57)$$

$$(1.5) \quad (162.0) \quad R^2 = 0.999,$$

حيث تظهر نسب t أسفل تقديرات المعاملات. تفسر هذه المعادلة، بوضوح، معظم التغير في الاستهلاك (كما يظهر من R^2 التي تقترب من الواحد الصحيح)، وأكثر من ذلك، فإن تباين \hat{b}_i صغير جداً، كما يستدل عليه من القيمة الهائلة لنسبة t المناظرة.

جدول رقم (١-٦) بيلالين الدولارات الأمريكية

السنة	الانفاق الاستهلاكي	الدخل المتاح
١٩٥١	٢٠٦,٣	٢٢٦,٦
١٩٥٢	٢١٦,٧	٢٣٨,٣
١٩٥٣	٢٣٠,٠	٢٥٢,٦
١٩٥٤	٢٣٦,٥	٢٥٧,٤
١٩٥٥	٢٥٤,٤	٢٧٥,٣
١٩٥٦	٢٦٦,٧	٢٩٣,٢
١٩٥٧	٢٨١,٤	٣٠٨,٥
١٩٥٨	٢٩٠,١	٣١٨,٨
١٩٥٩	٣١١,٢	٣٣٧,٣
١٩٦٠	٣٢٥,٢	٣٥٠,٠
١٩٦١	٣٣٥,٢	٣٦٤,٤
١٩٦٢	٣٥٥,١	٣٨٥,٥
١٩٦٣	٣٧٥,٠	٤٠٤,٦
١٩٦٤	٤٠١,٢	٤٣٨,١
١٩٦٥	٤٣٢,٨	٤٧٣,٢
١٩٦٦	٤٦٦,٣	٥١١,٩
١٩٦٧	٤٩٢,١	٥٤٦,٣
١٩٦٨	٥٣٦,٢	٥٩١,٠
١٩٦٩	٥٧٩,٦	٦٣٤,٢

المصدر: التقرير الاقتصادي الرئيسي، واشنطن، يناير ١٩٧٢، صفحة ٢١٢.

دعنا الآن نفحص القيم المقدرة للأخطاء العشوائية لنرى ما إذا كانت هناك أي إشارة لوجود الارتباط الذاتي. وتحسب معظم برامج الحاسوب لتحليل الانحدار قيمة إحصائية d لديربان واتسون، ولذلك؛ فإن التساؤل يمكن الإجابة عليه مباشرة. ولكن لتتعرف على هذا المنهج فإن الجدول رقم (٦-٢) يظهر التسلسل الفعلي للحسابات. ولحساب الأخطاء العشوائية، نستخدم أولاً المعادلة المقدرة (6.57) لحساب القيمة المتوقعة للاستهلاك كل سنة ثم نطرح هذه القيمة من الاستهلاك الفعلي للحصول على تقديرات الأخطاء العشوائية التي تظهر في العمود الرابع. فإذا نظرت إلى الأرقام في هذا العمود فسوف تلاحظ أن هناك تعاقبات للأخطاء العشوائية المقدرة، فقيم سالبة تتبعها سلسلة من القيم الموجبة للأخطاء وهذه تجعلنا نشك فوراً لأن ذلك يوحي بوجود ارتباط متسلسل موجب بين الأخطاء العشوائية وتجعلنا نتوقع قيمة منخفضة للإحصائية d . وفي الحقيقة فإن قيمة الإحصائية d منخفضة جداً عن ٢ حيث تعادل ١,٠١، فإذا مارجعنا مرة ثانية للجدول الإحصائي ٤ نجد أنه وحسب اختبار ديربن واتسون للإحصائية d (الاختبار ذو الطرفين)*. أنه إذا كانت $\alpha = 0.5, k=1, n=19$ فإن قيمة الحد الأدنى، هي 1.06، وإحصائيا تقع d أسفل هذا الحد الأدنى. ولذلك نرفض فرضية العدم بعدم وجود ارتباط ذاتي لصالح الفرضية البديلة بوجود الارتباط الذاتي بين الأخطاء العشوائية ($\gamma \neq 0$).

ولاستخدام المنهج الموضح سابقاً لتصحيح النموذج من الارتباط الذاتي، ينبغي أن نقدر أولاً العلاقة بين الأخطاء العشوائية. نفترض أن هذه العلاقة تأخذ الشكل التالي:

$$u_t = \gamma u_{t-1} + \varepsilon_t$$

حيث إن ε_t تحقق الافتراضات كافة التي وضعناها من قبل، وبأخذ القيم المقدرة للأخطاء العشوائية في الجدول رقم (٦-٢) نستخدم الصيغة الموجودة في (6.45) لتقدير قيمة γ :

* يجب تكوين الافتراضات قبل اختبار النتائج، ولهذا السبب نستخدم الاختبار ذو الطرفين، ويتضمن ذلك أنه قبل اختبار البواقي فعلياً، لا يكون لدينا سبب مقنع لتوقع أنه إذا كانت لدينا مشكلة ارتباط متسلسل ذاتي فإن ذلك الارتباط سيكون موجباً.

جدول (٧-٩)

Year	Actual Consumption (C_t)	Predicted Consumption (\hat{C}_t)	$\hat{u}_t = C_t - \hat{C}_t$	\hat{u}_t^2	\hat{u}_{t-1}	$\hat{u}_t - \hat{u}_{t-1}$	$(\hat{u}_t - \hat{u}_{t-1})^2$
1951	206.3	208.6	-2.3	5.2	-2.3	-0.2	0.0
1952	216.7	219.2	-2.5	6.2	-2.5	0.4	0.1
1953	230.0	232.1	-2.1	4.6	-2.1	2.1	4.6
1954	236.5	236.5	0	0	0	1.7	2.8
1955	254.4	252.7	1.7	2.9	1.7	-3.9	15.3
1956	266.7	268.9	-2.2	5.0	-2.2	0.8	0.7
1957	281.4	282.8	-1.4	1.9	-1.4	-0.6	0.4
1958	290.1	292.1	-2.0	4.1	-2.0	4.3	18.8
1959	311.2	308.9	2.3	5.4	2.3	2.5	6.2
1960	325.2	320.4	4.8	23.1	4.8	-3.0	9.3
1961	335.2	333.4	1.8	3.1	1.8	1.0	0.9
1962	355.1	352.4	2.7	7.4	2.7	2.4	5.8
1963	375.0	369.9	5.1	26.5	5.1	-4.2	17.2
1964	401.2	400.2	1.0	1.0	1.0	-0.2	0.0
1965	432.8	432.0	0.8	0.6	0.8	-1.6	2.4
1966	466.3	467.1	-0.8	0.6	-0.8	-5.4	28.8
1967	492.1	498.2	-6.1	36.7	-6.1	3.6	13.0
1968	536.2	538.7	-2.5	6.4	-2.5	4.3	18.2
1969	579.6	577.9	1.7	3.0	1.7		
				$\sum \hat{u}_t^2 = 143.7$		$\sum (\hat{u}_t - \hat{u}_{t-1})^2 = 144.5$	
				$d = \frac{\sum (\hat{u}_t - \hat{u}_{t-1})^2}{\sum \hat{u}_t^2} = \frac{144.5}{143.7} = 1.01$			

^a Figures may not sum precisely due to rounding.

$$\hat{\gamma} = \frac{\sum_{t=2}^n \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^n \hat{u}_{t-1}^2} = 0.48. \quad (6.58)$$

وباستخدام قيمة مقدرة لـ γ مساوية لـ 0.48، نحسب بعد ذلك:

$$C_t^* = C_t - \hat{\gamma} C_{t-1} = C_t - 0.48 C_{t-1},$$

$$Y_{dt}^* = T_{dt} - \hat{\gamma} Y_{d(t-1)} = Y_{dt} - 0.48 Y_{d(t-1)}.$$

وبتكوين انحدار لـ C^* على Y_d^* نجد أن:

$$\hat{C}_t^* = 2.12 + 0.905 Y_{dt}^*, \quad n=18, \quad (6.59)$$

$$(1.0) \quad (98.9) \quad R^2 = 0.998.$$

وأخيرا فإن تقديراتنا للحد الثابت ولتباينه هي:

$$\hat{b}_0 = \frac{\hat{B}}{1 - \hat{\gamma}} = \frac{2.12}{(1 - 0.48)} = 4.08$$

$$\text{var}(\hat{b}_0) = \frac{1}{(1 - \hat{\gamma})^2} \text{var}(\hat{B}) = \frac{1}{(1 - 0.48)^2} (4.2) = 15.5.$$

وتكون معادلتنا المقدرة والمصححة من الارتباط الذاتي هي*:

$$\hat{C}_t = 4.08 + 0.905 Y_{dt} \quad (1.0) \quad (98.9)$$

لاحظ أنه، على الرغم من أن القيمة المقدرة للميل الحدي للاستهلاك (م ح س) b_1 ، في هذه الحال، تأخذ فعليا قيمة معادلة الانحدار العادية نفسها (6.50) فإن نسبة t المناظرة أصبحت أقل بدرجة كبيرة عندما صححنا مشكلة الارتباط الذاتي، وفي حالات أخرى، قد يحسم هذا الخلاف بين رفض فرضية الغدم بوجود القيمة الصفرية للمعلمة أو قبولها.

* حسب نسبة t للحد الثابت عن طريق قسمة الانحراف المعياري لـ \hat{b}_0 (الجذر التربيعي لتباينه) على \hat{b}_0 .

الارتباط الذاتي والمتغيرات التابعة المبطة

يعد اختبار ديربن واتسون غير صحيح، في حال، احتواء نموذج الانحدار على قيمة مبطة للمتغير التابع باعتبارها واحدا من المتغيرات المستقلة. ولفهم المشاكل المتضمنة افترض النموذج التالي:

$$Y_t = b_0 + b_1 X_{1t} + \dots + b_k X_{kt} + aY_{t-1} + u_t \quad (6.60)$$

$$u_t = \gamma u_{t-1} + \varepsilon_t, \quad t = \dots, -2, -1, 0, 1, 2, \dots \quad (6.61)$$

حيث إن ε_t لها قيمة متوقعة صفرية $E(\varepsilon_t) = 0$ وتباين ثابت $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ و $E(\varepsilon_t \varepsilon_s) = 0$ إذا كانت $t \neq s$. نفترض، أيضا، أن $|\gamma| < 1$ وأن $|a| < 1$ يتماثل هذا النموذج مع النموذج الموجود في (6.50) باستثناء أنه في (6.60)، يحتوي على متغير مستقل Y_{t-1} بينما لا يحتوي (6.50) على ذلك المتغير. وقد افترضنا أن $|a| < 1$ للسبب نفسه الذي افترضنا من أجله $|\gamma| < 1$.*

لاحظ أنه، في مرحلتنا هذه، طالما أن Y_t يعتمد على u_t في (6.60) فإن Y_t و u_t مرتبطان. وبالمثل، طالما أن Y_{t-1} تعتمد، بدورها، على u_{t-1} (من خلال الصيغة المبطة لـ (6.60) فإنه يمكن إثبات أن هذين المتغيرين مرتبطان أيضا. وأخيرا، ينتج عن ذلك - بديهيا، في الأقل - أنه إذا كانت $\gamma \neq 0$ فإن Y_{t-1} تكون مرتبطة بـ u_t ، لأنه من خلال (6.61) نجد أن u_t تعتمد على u_{t-1} ، و Y_{t-1} مرتبطة بـ u_{t-1} .

إذا قمنا بتقدير (6.60) باستخدام منهج المتغير المساعد، فإن واحدة من المعادلات الطبيعية وهي $\sum_{t=1}^n (u_{t-1} \hat{u}_t) = 0$ سوف تناظر الافتراض $\text{cov}(Y_{t-1}, u_t) = 0$ ولكننا نلاحظ أنه إذا كانت $\gamma \neq 0$ فإن $\text{cov}(Y_{t-1}, u_t) \neq 0$ ، لذلك إذا ما استخدمنا $\sum (Y_{t-1} \hat{u}_t) = 0$ بوصفها إحدى هذه المعادلات الطبيعية، فإن مقدراتنا سوف تكون متحيزة. ويمكن إثبات أنها ستكون، أيضا، غير متسقة. والمشكلة الرئيسية في كل هذا هي أن ذلك التحيز وعدم الاتساق يجعل إحصائية d (لديران واتسون) قريبة من 2 حتى ولو كانت $\gamma \neq 0$. تذكر أن إحصائية d تقترب من 2 إذا كانت $\gamma \neq 0$. تذكر، أيضا، أن قيمة d القريبة من 2 تؤدي إلى قبول الافتراض

* نحتاج هذا الافتراض لجعل النموذج مستقراً. انظر ملحق هذا الفصل لفهم دور هذا الافتراض.

بعدم وجود ارتباط ذاتي ($\gamma = 0$). ودلالة كل هذا هي أن استخدام اختبار ديربن واتسون في حالات تحتوي على متغيرات تابعة متباطئة، سيؤدي بالباحث إلى قبول الافتراض بعدم وجود ارتباط ذاتي بغض النظر عن وجود ذلك الارتباط الذاتي أو عدم وجوده. ومن الواضح أنه ينبغي ألا يستخدم اختبار ديربن واتسون في هذه الحالات.

ولحسن الحظ، يوجد اختبار آخر للارتباط الذاتي في حالة وجود متغير تابع متباطئ. ويتضمن هذا الاختبار استخدام إحصائية h لديربن. * افترض أنه تم تقدير (6.60) بواسطة طريقة المتغير المساعد التقليدية، وأن γ قد قدرت بواسطة (6.45). دع $\hat{\gamma}$ و \hat{a} مقدرات لـ γ و a تم الحصول عليها، دع $\text{var}(\hat{a})$ أيضا، مقدرا لتباين \hat{a} وأنه قد حصل عليه، حينئذ فإن إحصائية h تكون:

$$h = \hat{\gamma} \left(\frac{n}{1 - n \text{var}(\hat{a})} \right)^{1/2} \quad (6.62)$$

حيث n حجم العينة، يمكن إثبات أنه إذا لم يكن هناك ارتباط ذاتي (أي $\gamma = 0$) فإن h سوف تكون موزعة توزيعا طبيعيا معياريا $N(0,1)$ إذا كان حجم العينة لانهائيا. أما إذا كان هناك ارتباط ذاتي فإن h تصبح كبيرة وبالتحديد إذا كانت $\gamma > 0$ فإن h ستصبح كبيرة في الاتجاه الموجب. وعلى العكس إذا كانت $\gamma < 0$ تصبح h كبيرة ولكن في الاتجاه السالب.

توحي الملاحظات السابقة بالاختبار التالي لوجود الارتباط الذاتي في حالة وجود المتغيرات التابعة المبطة بوصفها متغيرات مستقلة. هذا الاختبار هو اختبار للعينات الكبيرة، أو بمعنى آخر تكون النتائج صحيحة أو دقيقة، فقط، إذا كانت العينة ذات حجم لانهائي.

* يرجع إلى:

J. Durbin, "Testing for Serial Correlation in Least Squares Regression When Some of the Regressors are Lagged Dependent Variables," *Econometrica* 38 (1970), pp. 410-421.

اعتبر النموذج في (6.60)، وفرضية العدم $H_0: \gamma = 0$. دع الفرضية البديلة $H_1: \gamma \neq 0$. حينئذ، يمكننا اختبار H_0 مقابل H_1 عند مستوى معنوية 0.05 بالطريقة التالية:

١- قدر (6.60) بالطريقة العادية مستخدما طريقة المتغير المساعد من الفصل الرابع ولاحظ قيمة $\text{var}(\hat{\gamma})$ التي حصل عليها.

٢- من البواقي، تحسب $\hat{\gamma}$ كما في (6.45). وبالمقابل إذا كان البرنامج يزودنا بإحصائية d لديرين واتسون فإنه يمكننا استخدام التقريب $\hat{\gamma} = 1 - d/2$. ويبنى هذا التقريب على (6.54).

٣- نحسب قيمة إحصائية h لديرين.

٤- نرفض H_0 إذا كانت $|h| > 1.96$ ونقبلها إذا كانت h غير ذلك.

هناك تعديلات على هذا الاختبار (ينبغي أن يكونا واضحين). الأول إذا كان $\gamma > 0$ $H_1: \gamma > 0$ واحتفظنا بمستوى 0.05 باعتباره مستوى معنوية فإننا سنرفض H_0 إذا كانت $h > 1.645$ ، والثاني إذا كان $\gamma < 0$ $H_1: \gamma < 0$ فسنرفض H_0 إذا كانت $h < -1.645$ أما الاختبارات عند مستويات معنوية أخرى (مثلا عن 0.01) فهي تطبيقات مباشرة ونتركها للقارئ على سبيل التدريب.

هناك حدود على الاختبار المبني على إحصائية h ، فعلى سبيل المثال، وفي ضوء (6.62) ينبغي أن يكون واضحا أن الاختبار يفشل إذا:

$$n \text{var}(\hat{a}) \geq 1 \quad (6.63)$$

وذلك نظرا لأن h ستضمن الجذر التربيعي لرقم سالب. في مثل هذه الحالات لا يمكن تعريف h . لا تمثل هذه الحالة مشكلة من الناحية النظرية فحسب بل تحدث غالبا، في التطبيق، وفي مثل هذه الحالات، يمكن استخدام اختبار بديل. وهذا الاختبار هو اختبار للعينات الكبيرة، ولذا، تعد نتائج تقريبية من الناحية التطبيقية. مرة أخرى، دع فرضية العدم $H_0: \gamma = 0$ والفرضية البديلة $H_1: \gamma \neq 0$ ، بعدئذ، إذا أخذنا مستوى المعنوية عند 0.05، فإن خطوات الاختبار المقترح هي:

- ١- نقدر المعادلة الأساسية، مثلاً (6.60)، بطريقة المتغير المساعد التقليدية.
 - ٢- نحصل، بعد ذلك، على الأخطاء العشوائية المقدرة \hat{u}_t .
 - ٣- نقدر باستخدام النتائج التي توصل إليها من الخطوة ٢ معادلة الانحدار التالية:

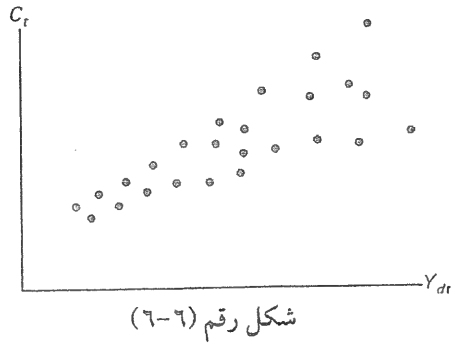
$$\hat{u}_t = a_0 + a_1 \hat{u}_{t-1} + a_2 Y_{t-1} + c_1 X_{1t} + \dots + c_k X_{kt} + \varepsilon_t \quad (6.64)$$
 بالطريقة العادية. لاحظ أن (6.64) تحتوي على الخطأ العشوائي المقدر باعتباره متغيراً مستقلاً وعلى كافة المتغيرات المستقلة في النموذج الأصلي (6.60).
 - ٤- نحصل على نسبة t المناظرة لـ a_1 (أي $\hat{a}_1 / \hat{\sigma}_{\hat{a}_1}$). يمكن إثبات أنه $\hat{a}_1 / \hat{\sigma}_{\hat{a}_1}$ سيكون لها توزيع $N(0,1)$ إذا كانت $\gamma = 0$ و n لانهائية.
- وبناء على ذلك، يتضمن الاختبار رفض H_0 إذا كان $\hat{a}_1 / \hat{\sigma}_{\hat{a}_1} > 1.96$ وقبلها في الحالات الأخرى. فإذا كانت الفرضية البديلة هي $H_1: \gamma > 0$ فنرفض H_0 إذا كان $\hat{a}_1 / \hat{\sigma}_{\hat{a}_1} > 1.645$. وتركت الحال التي يكون فيها الفرض البديل $H_1: \gamma > 0$ على سبيل التدريب للقارئ.

(٦-٣) اختلاف التباين

نتناول في هذا المبحث مشكلة أخرى تنشأ نتيجة انتهاك أحد الافتراضات المرتبطة بالأخطاء العشوائية. تذكر أننا افترضنا في نموذجنا الأساسي للانحدار أن:

$$\text{var}(u_t) = E(u_t^2) = \sigma_u^2$$

أي أننا افترضنا أن جميع الأخطاء العشوائية لها التباين نفسه، وتعرف هذه الحالة (اصطلاحاً) بثبات التباين homoscedasticity. ولكن، قد يختلف تباين الأخطاء العشوائية، وفي هذه الحالة يطلق على هذه الحالة إختلاف التباين. وعلى سبيل المثال، قد نجد من دراسة لمستويات الإنفاق الاستهلاكي لأسر ذات دخول متاحة مختلفة، أن التباين في الاستهلاك يزداد مع ازدياد مستوى الدخل. فالأسر ذات الدخل المرتفع، مثلاً، قد تتميز بمرونة أكبر في الاستهلاك. ويظهر هذا الشرط بوضوح في الشكل رقم (٦-٦) حيث نجد أن مدى التغير للمجموعة الافتراضية من النقاط يتزايد عند مستويات الدخل الأعلى، وفي حالة مثل هذه، قد نفترض أن الخطأ العشوائي في دالة الاستهلاك يتسم باختلاف التباين.



نموذج أساسي

افترض أن لدينا دالة للاستهلاك تأخذ الشكل

$$C_t = b_0 + b_1 Y_{dt} + b_2 A_t + u_t, \quad (6.65)$$

حيث:

$$C_t = \text{الانفاق الاستهلاكي للأسرة } t,$$

$$Y_t = \text{الدخل المتاح للأسرة } t,$$

$$A_t = \text{الاصول السائلة التي تمتلكها الأسرة } t, \text{ وأخيراً}$$

$$u_t = \text{الخطأ العشوائي.}$$

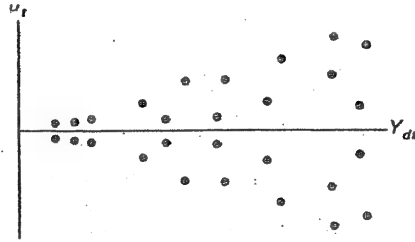
نفترض الآن أنه لأي مجموعة من قيم المتغيرات المستقلة، فإن u_t موزع توزيعاً طبيعياً، وغير مرتبط ذاتياً لكن تباينه يرتبط بشكل متناسب مع دخل الأسرة t ، أي $\text{var}(u_t) = Y_{dt} \sigma_{u_t}^2$. وهكذا، كلما كان الدخل كبيراً نتوقع أن نشاهد تبايناً أكبر في الاستهلاك.

ولقد افترضنا في نماذجنا السابقة أن الخطأ العشوائي مستقل عن جميع المتغيرات المستقلة، ولكن، ليس بوسعنا الآن أن نحفظ بهذا الافتراض، لأننا حددنا أن حجم التباين يعتمد على قيمة أحد المتغيرات المستقلة Y_{dt} . لذلك، لم يعد الخطأ العشوائي مستقلاً عن ذلك المتغير المستقل. وبدون وضع فروض إضافية، ليس بوسعنا أن نستخدم منهجنا العادي في التقدير، ذلك لأن التباين بين الخطأ العشوائي والمتغير المستقل Y_{dt} لا يساوي الصفر. وكما سنبين فيما

بعد، فإن الافتراض الإضافي الذي يمكننا من معالجة مشكلة اختلاف التباين ومن افتراض أن التباين بين الخطأ العشوائي وبين المتغير المستقل Y_{dt} يساوي الصفر، وكيفما كانت قيم المتغيرات المستقلة Y_{dt} و A_t لأي مشاهدة هو أن يكون المتوسط الحسابي للخطأ العشوائي مساوياً للصفر. واصطلاحاً فإنه بالنسبة لأي قيم معطاة لكل من A_s, Y_{ds} تكون $E(u_t) = 0$ لجميع t و s . وبالنسبة للمعادلة (6.65)، فهذا يدل على أن القيمة المتوقعة لـ C_t^m لاتزال هي:

$$C_t^m = b_0 + b_1 Y_{dt} + b_2 A_t$$

ويدل هذا الافتراض، أيضاً، على أن الخطأ العشوائي غير مرتبط بأي من المتغيرات المستقلة، أي $\text{cov}(u_t, Y_{dt}) = \text{cov}(u_t, A_t) = 0$. فإذا كان u_t مرتبطاً مثلاً بـ Y_{dt} فإنه يتوقع أن تزداد قيمته أو تنقص مع زيادة Y_{dt} . ولكن افتراضنا بأن القيمة المتوسطة لـ u_t تساوي الصفر لأي قيمة من قيم Y_{dt} يتضمن أن ذلك ليس صحيحاً، حيث إنه، مع زيادة Y_{dt} ، تظل القيمة المتوقعة لـ u_t ثابتة، أي صفراً. ينتج عن ذلك أن u_t و Y_{dt} غير مرتبطتين.

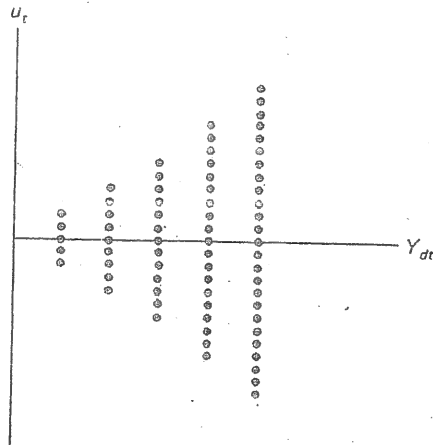


شكل رقم (٦-٧)

قد تبدو هذه النتيجة محيرة في ضوء افتراضنا بأن تباين u_t يزداد مع Y_{dt} . ولكن تزول هذه الحيرة بالنظر إلى الشكل (٦-٧) حيث يظهر من مجموعة النقاط الافتراضية أن تباين u_t يزداد مع تزايد Y_{dt} ، ولكن من الواضح، أيضاً، أن u_t لن يكون مرتبطاً مع Y_{dt} لأن القيمة المتوسطة لـ u_t تساوي الصفر عند أي قيمة لـ Y_{dt} .

لاحظنا مما سبق أن u_t و Y_{dt} (وكذلك u_t و A_t) ليستا مرتبطتين، لأن القيمة المتوسطة لـ u_t تساوي الصفر عند أي قيمة معطاة لـ Y_{dt} (وكذلك A_t). فإذا كنا قد افترضنا أن قيمة u_t المتوسطة مساوية للصفر بدون التصريح بالشرط «لكل قيمة معطاة لـ Y_{dt} » فإنه لا يمكننا الحصول على هذه النتيجة. افترض (مثلاً) أن متغير قيمته المتوسطة تساوي الصفر، $E(X_1) = 0$. دع $X_2 = 2X_1$ حينئذ تكون القيمة المتوسطة لـ X_2 مساوية للصفر $E(X_2) = 2E(X_1) = 0$ ، ولكن القيمة المتوسطة لـ X_2 لن تكون مساوية للصفر عند أي قيمة معطاة لـ X_1 . على سبيل المثال إذا كانت $X_1 = 3$ فإن القيمة المتوسطة لـ X_2 تكون 6. في هذه الحال يكون X_1 و X_2 مرتبطان ببعضهما بصورة تامة.

وقبل أن نتجه إلى مشاكل التقدير علينا أن نوضح نقطة أخيرة قد تكون غامضة لكثير من القراء. نرى في الشكل رقم (٦-٨) وهو شكل منح من الشكل رقم (٦-٧) أن جميع النقاط المناظرة لأي قيمة من Y_{dt} تظهر لها قيمة متوقعة مساوية للصفر. هذا يعكس أن القيمة المتوسطة لـ u_t تساوي الصفر لأي قيمة معطاة من Y_d .



شكل رقم (٦-٨)

لاحظ أن القيمة المتوسطة لجميع النقاط في الشكل تبدو مساوية للصفر، وينظر هذا شرط أن القيمة المتوسطة (يطلق عليها، في بعض الأحيان، «المتوسط العام overall mean» لـ u هي الصفر. يتضح لنا الآن أن افتراض أن القيمة المتوسطة لـ u المناظرة لأي قيمة معطاة لـ Y_d المساوية للصفر تتضمن، بدورها، أن القيمة المتوسطة العامة مساوية للصفر أيضا. ولكن العكس ليس صحيحا بالضرورة، فقد تكون القيمة المتوسطة لـ u سالبة لبعضها قيم Y_d وموجبة لبعضها الآخر، ومع ذلك تظل القيمة المتوسطة العامة لـ u مساوية للصفر.

تأثيره على مقدراتنا

ماذا يحدث إذا استخدمنا المنهج العادي في التقدير لمعادلة تعاني مشكلة اختلاف التباين؟ بديها، يمكننا أن نتخيل أنه طالما أن القيمة المتوسطة للخطأ العشوائي لاتزال تساوي الصفر، وطالما أن u_t لاتزال غير مرتبطة بكل متغير من المتغيرات المستقلة [لذلك يكون $E(u_t Y_{dt}) = 0$ و $E(u_t A_t) = 0$ في المعادلة (6.65) فإن مقدرات معلمتنا سوف تظل متسقة وغير متحيزة.* وأساسا فالنقطة المهمة هي أن شروطنا $E(u_t) = 0$ ، $E(u_t Y_{dt}) = 0$ و $E(u_t A_t) = 0$ لاتزال تفيد بأن $\Sigma \hat{u}_t = 0$ و $\Sigma(\hat{u}_t Y_{dt}) = 0$ وأخيرا $\Sigma(\hat{u}_t A_t) = 0$. فأساسا لا يوجد خطأ في المعادلات الطبيعية. ولكن، كما هو الحال عند وجود الارتباط الذاتي، تكون هناك صيغ مختلفة لتباينات مقدرات معلمتنا. ومرة أخرى، إذا ما استخدمنا صيغنا العادية لتقدير هذه التباينات فستكون اختبارات الفرضيات وتكوين فترات الثقة الناتجة مشكوك فيها. ينبغي أن يكون هذا واضحا، طالما أن نموذجنا الأساسي للانحدار يفترض

* يمكن للقارئ المهتم أن يثبت أن مقدراتنا المعتادة لاتزال غير متحيزة عن طريق العمل من خلال المناقشة الموجودة في ملحق الفصل الرابع. وعند القيام بذلك، علينا أن نلاحظ أن الافتراض الوحيد الذي نحتاجه في عملية الاشتقاق هو افتراض أن القيمة المتوسطة للأخطاء العشوائية هي الصفر لأي قيمة من قيم المتغيرات المستقلة وينتج هذا الشرط - في النموذج المعتاد - من افتراض الاستقلال، وقد أخذنا بهذا الافتراض، صراحة، في نموذجنا الذي يعاني اختلاف التباين.

تباينا ثابتا، وأن منهجنا في التقدير ينتج مقدرًا لهذا الثابت. لكن، مع وجود اختلاف التباين، فلن يظل تباين الخطأ العشوائي ثابتًا، وإنما سيكون متغيرًا. وهذا يعني أن مقدرنا المعتاد سيمثل في الحقيقة أحد أنواع المتوسطات للتباينات المختلفة للأخطاء العشوائية، مثل هذا المقدر تكون له أهمية محدودة، ولا يسمح لنا - على سبيل المثال - ببناء فترات ثقة صحيحة (أو نسب t) لمعاملات المعادلة. وكما هو الحال عند وجود الارتباط الذاتي يمكننا الحصول على مقدرات أفضل لمعاملاتنا (أي مقدرات لها تباينات أصغر) ويمكننا أن نوجد مقدرات لهذه التباينات عن طريق إدخال معلومات ترتبط بالخصائص الحقيقية للخطأ العشوائي في منهجنا للتقدير.

طريقة للتقدير

لفحص مشكلة اختلاف التباين بعمق، دعنا نعود إلى علاقة الاستهلاك في المعادلة (6.65)، حيث جعلنا $\sigma_u^2 = Y_{dt}$. سنثبت الآن أنه إذا ماقسمنا حدود المعادلة (6.65) بوساطة $\sqrt{Y_{dt}}$ فإننا سنحصل على معادلة تتسم بثبات التباين للخطأ العشوائي وينتج عن تلك القسمة:

$$\frac{Ct}{\sqrt{Y_{dt}}} = b_0 \left(\frac{1}{\sqrt{Y_{dt}}} \right) + b_1 \sqrt{Y_{dt}} + b_2 \left(\frac{A_t}{\sqrt{Y_{dt}}} \right) + u_t^*, \quad (6.66)$$

$$u_t^* = \frac{u_t}{\sqrt{Y_{dt}}} \quad \text{حيث:}$$

ولما كانت القيمة المتوسطة لـ u_t هي الصفر لأي مستوى من مستويات Y_{dt} فسيكون لدينا: *

$$E(u_t^*) E \left(\frac{u_t}{\sqrt{Y_{dt}}} \right) = \left(\frac{1}{\sqrt{Y_{dt}}} \right) E(u_t) = 0. \quad (6.67)$$

أما بالنسبة لتباين u_t^* فسيكون لدينا وفقا للافتراضات نفسها:

* ينبغي أن يكون معنى (6.67) واضحًا. إذا أعطيت Y_{dt} على أنها 900، وأن $E(u_t^*) = E(u_t)/30 = 0$ طلالا أن

القيمة المتوسطة لـ u_t تساوي الصفر لأي قيمة من قيم Y_{dt} .

$$\begin{aligned} \text{var}(u_t^*) &= E(u_t^*)^2 = E\left(\frac{u_t^2}{Y_{dt}}\right) = \frac{1}{Y_{dt}} E(u_t^2) \\ &= \frac{1}{Y_{dt}} (Y_{dt}) \sigma_u^2 = \sigma_u^2. \end{aligned} \quad (6.68)$$

لنجد أن نموذجنا المعدل (6.66) هو نموذج يتسم فيه الخطأ العشوائي، u_t^* بقيمة متوقعة صفرية وتباين ثابت.

وبالاستمرار في تحليلنا يتبين لنا أنه ليس من الصعب اكتشاف أن u_t^* غير مرتبطة بالمتغيرات المستقلة في (6.66)، فعلى سبيل المثال، بالنسبة لأي مجموعة من المتغيرات المستقلة في (6.66) أو على نحو آخر، لأي قيم معطاة للمتغيرات المستقلة يكون $E(u_t^*) = (1/\sqrt{Y_{dt}}) E(u_t) = 0$. يترتب على النتائج السابقة أن u_t^* غير مرتبط بالمتغيرات المستقلة في (6.66) ولذلك نجد:

$$(1) \quad E\left[u_t^* \left(\frac{1}{\sqrt{Y_{dt}}}\right)\right] = 0,$$

$$(2) \quad E(u_t^* \sqrt{Y_{dt}}) = 0,$$

$$(3) \quad E\left[u_t^* \left(\frac{A_t}{\sqrt{Y_{dt}}}\right)\right] = 0.$$

باختصار، إذا قسمنا بالنسبة لكل مشاهدة كل من C_t ، Y_{dt} و A_t على $\sqrt{Y_{dt}}$ فإن نموذج الانحدار المناظر لهذه المجموعة الجديدة من القيمة الملاحظة «المصححة» سيصبح (6.66)، ويحقق هذا النموذج جميع افتراضاتنا الأساسية. وحيث يمكننا ببساطة استخدام طريقتنا المعتادة في التقدير للحصول، في هذه الحالة، على مقدرات غير متحيزة لكل من معاملات الانحدار وتباينات المقدرات،* وخاصة

* على العكس من النماذج السابقة، لا تحتوي المعادلة (6.66) (التي تناظر المعادلات الطبيعية في 6.69) على حد ثابت، ويترتب على ذلك تغير في صيغ التباينات لمقدراتنا. وبالتحديد، فإن حدود في هذه الصيغ سوف تعرف كبواقي انحدار (ولا تحتوي هذه على حد ثابت) المتغير المستقل رقم 1 على المتغيرات المستقلة الأخرى. على سبيل المثال، فإن تباين b_2 في (6.69) سيكون $\sigma_u^2 / \Sigma \hat{v}_{21}^2$ حيث إن \hat{v}_{21} هو المتبقي من الانحدار:

$$\frac{A_t}{\sqrt{Y_{dt}}} = \gamma_1 \left(\frac{1}{\sqrt{Y_{dt}}}\right) + \gamma_2 (\sqrt{Y_{dt}}) + v_{2t}$$

فإنه باستخدام الافتراضات من (1) إلى (3) تكون معادلاتنا الطبيعية هي*

$$\sum \left(\frac{C_t}{Y_{dt}} \right) = \hat{b}_0 \sum \left(\frac{1}{Y_{dt}} \right) + \hat{b}_1 n + \hat{b}_2 \sum \left(\frac{A_t}{Y_{dt}} \right),$$

$$\sum C_t = n \hat{b}_0 + \hat{b}_1 \sum Y_{dt} + \hat{b}_2 \sum A_t, \quad (6.69)$$

$$\sum \left(\frac{C_t A_t}{Y_{dt}} \right) = \hat{b}_0 \sum \left(\frac{A_t}{Y_{dt}} \right) + \hat{b}_1 \sum A_t + \hat{b}_2 \sum \left(\frac{A_t^2}{Y_{dt}} \right).$$

ويزودنا المثال الذي شرحناه الآن، أيضا، بنظرة أعمق لمشكلة اختلاف التباين، فمنهجنا العادي للتقدير يعطي وزنا متساويا لكل مشاهدة من المشاهدات عند ايجاد مقدرات معلمتنا. بينما توضيح مناقشتنا هذه أنه عند وجود مشكلة اختلاف التباين، فإنه ينبغي أن نعطي أوزانا مختلفة للمشاهدات، وبالتحديد، وعلى سبيل المثال ينبغي أن نعطي وزنا لكل مشاهدة يعادل $(1/\sqrt{Y_{dt}})$ ويعني ذلك أنه ينبغي أن نعطي وزنا أقل للمشاهدات المناظرة للتباينات الكبيرة عن تلك المناظرة لتباينات أقل، وبديهيًا يبدو هذا الأمر ذو معنى. فالمشاهدة التي تناظر تباينا أقل سوف تكون على الأرجح قريبة من خط الانحدار الحقيقي،** وفي منهجنا للتقدير ينبغي أن نعطي بعض الاهتمام لتلك النقاط التي نعتقد بأنها تقع أقرب إلى خط الانحدار الحقيقي عن تلك التي تكون في المتوسط بعيدة عنه، فالمشاهدات التي تناظر تباينات أقل هي ببساطة أكثر قيمة في تقدير موقع خط الانحدار عن تلك التي تكون انحرافاتهما عن الخط أكبر.

* ينبغي أن نشير هنا إلى أن المعادلات الطبيعية آنفة الذكر قد اشتقت عن طريق وضع الافتراضات (1)-(3). وفي هذه الحالة، فإن الافتراض $E(u_t) = 0$ لم يستخدم. أي أنه على الرغم من وجود ثلاث معلمات فقط، b_0 ، b_1 ، و b_2 فإنه يكون لدينا افتراضات أربعة خاصة بالخطأ العشوائي. وتتشأ هذه المشكلة، التي تتمثل في وجود عدد كبير من الافتراضات عادة، عند حل مشكلة اختلاف التباين والحل هو (عموماً) أن يتم القيام بنفس العمل الذي قمنا به في المتن، وهو أن نستخدم فقط تلك الافتراضات التي تناظر التغيرات للمتغيرات المستقلة في النموذج المعدل. ويتضمن ذلك أننا يجب أن نهمل الافتراض بأن القيمة المتوسطة للخطأ العشوائي في النموذج المعدل يساوي الصفر، وعلى الرغم من أن إثبات ذلك خارج عن نطاق هذا الكتاب، فإنه يمكننا أن نبين أنه إذا اتبعت الطريقة نفسها، فإن المقدرات الناتجة لـ b_0 ، b_1 و b_2 ستكون لها تباينات أصغر عما لو فرضنا القيمة المتوسطة المساوية للصفر واثنتين (أي اثنتين) من الافتراضات الثلاثة السابقة.

** نعتي بخط الانحدار الحقيقي معادلة القيمة المتوسطة التي تربط المتغير التابع بالمتغيرات المستقلة.

اختلاف التباين: طرق إضافية للمعالجة

كيف تعرف أنه لديك مشكلة اختلاف التباين؟ وكيف تغير - عموماً - من طريقتك في التقدير لمواجهتها؟ ليست هذه أسئلة يسهل حلها (ولسوء الحظ غالباً ما يتم تجاهلها). إحدى الطرق المقنعة لاكتشاف المشكلة هو أن نختبر أولاً العلاقة محل الدراسة لنرى ما إذا كان هناك أي سبب للاعتقاد بأن الخطأ العشوائي يتسم باختلاف التباين. وغالباً ماكتشف مشكلة اختلاف التباين من تكوين النموذج ذاته. افترض، على سبيل المثال، أننا نهتم باختبار الفرضية بأن مستوى الأرباح π بالدولارات يعتمد على حجم مؤسسة الأعمال (مشاراً إليه بوساطة قيمة أصولها A). يمكننا أن نعبر عن هذه العلاقة على النحو:

$$\pi_t = b_0 + b_1 A_t + u_t, \quad (6.70)$$

حيث تتوافر لدينا مشاهدات عن π و A لعدد n من مؤسسات الأعمال. في ظل هذا النموذج، يصعب علينا قبول أن الخطأ العشوائي له التباين نفسه. وبالتأكيد، فإن التباين في الأرباح سيكون أكبر بين منشآت مثل جنرال موتورز General Motors وستاندرد أويل Standard Oil عن المتاجر المحلية لبيع المنتجات الغذائية أو الأجهزة المنزلية، ويرجع هذا، فقط، إلى الاختلاف الكبير في الحجم المطلق لأرباحها. والنقطة المهمة هنا هي أننا نتوقع وجود علاقة طردية قوية بين قيمة A_t وتباين u_t . وبهذه المناسبة قد يكون هناك بعض المتغيرات المستقلة الأخرى في المعادلة (6.70) ولكن مشكلة اختلاف التباين تركز عادة على العلاقة بين واحد من المتغيرات المستقلة، وتباين الخطأ العشوائي، وعلى أي حال، في حالة مثل هذه، فإن رجحان وجود اختلاف التباين يكون كبيراً جداً.

توجد طريقتان أساسيتان يمكننا أن نستخدم أيًا منهما لحل مشكلة اختلاف التباين. أولاً: يمكننا إعادة صياغة العلاقة بطريقة يمكن معها إزالة اختلاف التباين. أي أنه يمكننا النظر إلى مشكلة اختلاف التباين على أنها مشكلة صياغة غير جيدة للنموذج، وهنا يمكننا أن نحل المشكلة بطريقة كفء عن طريق بناء نموذج أفضل. وفي ضوء مثالنا أعلاه، فقد يكون من الأفضل أن نختبر العلاقة بين معدل الأرباح، أي $\pi^* = \pi/A$ وحجم المنشأة بدلاً من العلاقة بين π و A. وحيث، يمكننا تقدير المعادلة:

$$\pi_{it}^* = b_0 + b_1 A_{it} + u_{it}, \quad (6.71)$$

وبدون وجود سبب معين لتوقع اختلاف معدلات الأرباح اختلافا كبيرا بين المنشآت الكبيرة أو الصغيرة، يمكننا، بثقة أكبر، أن نفرض تباينا ثابتا بين الأخطاء العشوائية. ثانياً: يمكننا أن نحاول تحديد نمط اختلاف التباين ودمج هذه المعلومة في منهجنا للتقدير، فمثلا يمكننا في حالة دالة الاستهلاك (6.65) التي اخترناها سابقاً أن نفترض أننا نعرف هذا النمط: يتناسب تباين u_{it} مع مستوى الدخل. وقد عالجنا هذه المشكلة عن طريق قسمة الحدود في نموذج الانحدار بوساطة الجذر التربيعي للدخل. ولكن، في كثير من الحالات، قد لا يكون لدينا سبب كاف لافتراض نمط معين. ففي نموذج الأرباح - الأصول (6.70)، على سبيل المثال، هل يكون تباين الخطأ العشوائي متناسباً مع مستوى الأصول A_{it} أم مع A_{it}^2 أو مع أي دالة أخرى لـ A_{it} . قد لانعرف الاجابة مسبقاً عن هذه الأسئلة.

إن تحديد نمط اختلاف التباين يعد مشكلة صعبة. وبعض الحلول المقترحة لهذه المشكلة خارج نطاق هذا الكتاب*. ولحسن الحظ، يوجد منهج مباشر يمكن أن يؤدي إلى نتائج مشجعة. ففي ظل توافر افتراضات معينة يمكننا هذا المنهج من اختبار وجود اختلاف التباين، وأيضاً من تحديد نمطه وأكثر من ذلك، فإن هذا المنهج يعتمد على المادة العلمية التي عالجناها في الفصول السابقة، ويسهل هذا، بالطبع، من فهمه، كما يزودنا ذلك أيضاً بمراجعة مفيدة للمادة العلمية السابقة. افترض أننا نرغب في تقدير العلاقة:

$$Y_{it} = b_0 + b_1 X_{1it} + b_2 X_{2it} + u_{it}, \quad (6.72)$$

وأننا نشك أن تباين u_{it} مرتبط بانتظام بقيمة X_{2it} مثلاً، أي أننا نعتقد أن:

$$\sigma_{u_{it}}^2 = f(X_{2it}). \quad (6.73)$$

فإذا عرفنا الشكل المحدد للدالة $f(X_{2it})$ ، فإنه يمكننا (كما سبق) أن نحل مشكلة اختلاف التباين ببساطة من خلال قسمة معادلتنا (6.72) بوساطة $\sqrt{f(X_{2it})}$ ، لأن تباين الخطأ العشوائي الناتج $u_{it}^* = u_{it} / \sqrt{f(X_{2it})}$ سوف يكون ثابتاً ويساوي الواحد

الصحيح.* وسوف تحقق المعادلة الجديدة افتراض ثبات التباين ومن ثم، يمكننا إكمال التحليل على النحو المعتاد.

والمشكلة التي نواجهها هي أن الدالة $f(X_{2t})$ ليست معلومة. وسيكون طريقنا للحل أن نقول أولاً بتقريب وتقدير $f(X_{2t})$ وبعد ذلك كما افترضنا من قبل - نقوم بقسمة (6.72) بوساطة الجذر التربيعي لمقدرنا $f(X_{2t})$. حيثئذ يمكننا أن نستخدم معادلة الانحدار المعدلة لاشتقاق مجموعة من المعادلات الطبيعية التي يمكن أن نحلها للحصول على مقدرات معلمتنا.

لاحظ أولاً أن افترضنا بوجود اختلاف التباين في (6.75)، يتضمن أنه لقيمة معينة لـ X_{2t} ، أن:

$$E(u_t^2) = f(X_{2t}). \quad (6.74)$$

افتراض الآن:

$$\varepsilon_t = u_t^2 - f(X_{2t}) \quad (6.75)$$

ولاحظ أنه بالنسبة لقيمة معينة لـ X_{2t} :

$$\begin{aligned} E(\varepsilon_t) &= E(u_t^2) - f(X_{2t}) \\ &= f(X_{2t}) - f(X_{2t}) = 0. \end{aligned}$$

أي أن القيمة المتوسطة لـ ε_t هي الصفر.

دعنا الآن نحل (6.75) للحصول على u_t^2 :

$$u_t^2 = f(X_{2t}) + \varepsilon_t. \quad (6.76)$$

لاحظ أن تفسير (6.76) واضح ومباشر، حيث تم التعبير عن u_t^2 كمجموع قيمته المتوسطة $f(X_{2t})$ ومتغير آخر ε_t يعكس انحرافه عن قيمته المتوسطة. لاحظ، أيضاً، أن (6.76) يشبه كثيراً نموذج الانحدار.

* لاحظ أنه بالنسبة لقيمة معطاة من فإن:

$$E(u_t^*) = E\left[\frac{u_t^2}{f(X_{2t})}\right] = \frac{1}{f(X_{2t})} E(u_t^2) = \frac{f(X_{2t})}{f(X_{2t})} = 1.$$

دعنا نفترض الآن أنه توجد لدينا مشاهدات عن u_t ، وأن معلوماتنا المسبقة حول العلاقة تفيد أنه إذا كانت u_t تتسم باختلاف التباين كما في (6.74) فإن الدالة $f(X_{2t})$ قد يمكن، إلى حد ما، تقريبها بوساطة متعدد الحدود من الدرجة k . * يمكننا تحويل (6.76) في ظل هذه الافتراضات إلى نموذج انحدار يأخذ الشكل النمطي:

$$u_t^2 = a_0 + a_1 X_{2t} + \dots + a_k X_{2t}^k + \varepsilon_t. \quad (6.77)$$

لاحظنا أنه بالنسبة لأي قيمة معطاة من X_{2t} تكون $E(\varepsilon_t)$ ، كما لاحظنا أيضا أن هذا الشرط (القيمة المتوسطة الصفرية) تتضمن أن ε_t غير مرتبطة بـ X_{2t} . ولما كانت القيمة المعطاة لـ X_{2t} تتضمن أيضا، قيمة معطاة لكل قوى X_{2t} ، فإنه ينتج عن ذلك أن ε_t يكون غير مرتبط أيضا، بكل واحدة من هذه القوى X_{2t} . ** وهذا يعني أنه يمكننا أن نعامل (6.77) كنموذج انحدار له خطأ عشوائي ε_t يحقق الشروط كافة التي يمكن بوساطتها اشتقاق المعادلات الطبيعية.

افترض أننا نقوم بتقدير (6.77) بطريقتنا النمطية عن طريق جعل:

$$\sum \hat{\varepsilon}_t = 0, \sum (\hat{\varepsilon}_t X_{2t}) = 0, \dots, \sum (\hat{\varepsilon}_t X_{2t}^k) = 0.$$

حينئذ يمكننا أن نثبت (باستخدام بعض الافتراضات الإضافية) - أن المقدرات الناتجة $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_k$ متسقة. وهكذا يكون المقدر المتسق لـ $f(X_{2t})$:

$$\hat{f}(X_{2t}) = \hat{a}_0 + \hat{a}_1 X_{2t} + \dots + \hat{a}_k X_{2t}^k. \quad (6.78)$$

والآن ينبغي أن يكون باقي هذا المنهج واضحا. حيث سنقسم نموذجنا الأولي

$$(6.72) \text{ بوساطة } \hat{f}_t = \left[\hat{f}(X_{2t}) \right]^{\frac{1}{2}}, \text{ وبعثذ، نحصل على مقدراتنا لـ } b_2, b_1, b_0 \text{ من}$$

المعادلات الطبيعية:

* عادة تعد $k \geq 3$. وتعد نتائج هذا البحث صحيحة، فقط، على المستوى النظري من التحليل إذا كان

تقريب متعدد الحدود تاما. ولما كان يندر وقوع ذلك عمليا، فينبغي أن تعد جميع هذه النتائج تقريبية.

** على سبيل المثال، إذا كانت القيمة المتوسطة لـ ε_t هي الصفر على افتراض أن $X_{2t} = 3$ فإنها تكون مساويا

الصفر أيضا عند $X_{2t}^2 = 9$.

$$\begin{aligned}\sum \left(\frac{Y_t}{\hat{f}_t} \right) &= \hat{b}_0 \sum \left(\frac{1}{\hat{f}_t} \right) + \hat{b}_1 \sum \left(\frac{X_{1t}}{\hat{f}_t} \right) + \hat{b}_2 \sum \left(\frac{X_{2t}}{\hat{f}_t} \right), \\ \sum \left(\frac{Y_t X_{1t}}{\hat{f}_t} \right) &= \hat{b}_0 \sum \left(\frac{X_{1t}}{\hat{f}_t} \right) + \hat{b}_1 \sum \left(\frac{X_{1t}^2}{\hat{f}_t} \right) + \hat{b}_2 \sum \left(\frac{X_{1t} X_{2t}}{\hat{f}_t} \right), \\ \sum \left(\frac{Y_t X_{2t}}{\hat{f}_t} \right) &= \hat{b}_0 \sum \left(\frac{X_{2t}}{\hat{f}_t} \right) + \hat{b}_1 \sum \left(\frac{X_{1t} X_{2t}}{\hat{f}_t} \right) + \hat{b}_2 \sum \left(\frac{X_{2t}^2}{\hat{f}_t} \right).\end{aligned}$$

ولأن \hat{f}_t مقدر متسق لـ f_t ، فإنه يمكن إثبات (في ظل تحقق افتراضاتنا) أن المقدرات الناتجة تكون متسقة وكفءاً. يضاف إلى ذلك أنه إذا كان حجم العينة لانهائياً، فإن صيغنا العادية للتباين تكون صحيحة. دع $\hat{\sigma}_{\hat{b}_i}^2$ حيث $i=0,1,2$ تشير إلى مقدر التباين \hat{b}_i والذي نحصل عليه بواسطة صيغتنا العادية للتباين. حيث إن اختبارات الفرضيات وتكوين فترات الثقة يمكن أن يتم انشاؤها عن طريق فرض أن $(\hat{b}_i - b_i) / \hat{\sigma}_{\hat{b}_i}^2$ متغير طبيعي معياري. وتكون هذه النتيجة (مرة أخرى) صحيحة، فقط، في حالة العينة اللانهائية. ولذلك ينبغي علينا عند التطبيق أن ننظر إلى النتائج على أنها تقريبية وبطريقة مشابهة لحالة نموذج الارتباط الذاتي يكون سبب هذا التعقيد هو أن مقدر \hat{b}_i غير خطي في الخطأ العشوائي بسبب اعتماده على \hat{f}_t .

والصعوبة الواضحة في المنهج السابق هي أنه لن تتوافر لدينا مشاهدات عن u_t^2 ولذا، وقبل استخدام هذا المنهج، ينبغي علينا أولاً أن نقدر قيم u_t^2 . ويمكننا القيام بذلك بسهولة لأنه يمكننا أن نحصل على مقدرات متسقة للمعاملات، ومن ثم، للأخطاء العشوائية لنموذجنا الأصلي الذي يتسم باختلاف التباين (6.72) بواسطة طرقنا المعتادة في التقدير. حيث نقدر ببساطة معاملات النموذج (6.72) بالطرق المعتادة ثم نحصل بعد ذلك على مقدرنا للأخطاء العشوائية $\hat{u}_t = Y - \hat{b}_0 - \hat{b}_1 X_{1t} - \hat{b}_2 X_{2t}$ ويمكن إثبات أنه إذا نفذ المنهج السابق بعد إحلال \hat{u}_t^2 محل u_t^2 فإن النتائج السابقة تظل صحيحة كافة.

اختبار لاختلاف التباين

تتوافر لدينا طريقة لتصحيح طريقتنا في التقدير في ظل وجود مشكلة اختلاف التباين للأخطاء العشوائية. ولإكمال هذا المبحث، دعنا نعود إلى قضية تحديد ما إذا كان نموذج الانحدار يعاني مشكلة اختلاف التباين أم لا. لقد ناقشنا من قبل (في نموذجنا المفترض لانحدار أرباح المنشآت على أصولها) كيف يمكننا أن نختبر معادلة الانحدار لمعرفة ما إذا كان تكوين النموذج نفسه يوصي برجحان وجود اختلاف التباين، ولكن، من المرغوب فيه، وجود منهج نظري لاكتشاف هل يوجد اختلاف التباين للأخطاء العشوائية. سوف نقدم الآن مثل هذا الاختبار. والفرضية التي سنقوم باختبارها هي أن الأخطاء العشوائية في نموذجنا الأولى (6.72) تتسم باختلاف التباين من النوع الموجود في (6.73). ينبغي علينا أن نعرف أن (6.73) تعني، رياضياً، أنه «إذا كانت الأخطاء العشوائية» u_t تتسم باختلاف التباين، فإن نوع اختلاف تباينها سيكون كالمعطي في (6.73).

سوف نستخدم الآن تقريبنا لمتعدد الحدود (X_{2t}) لتكوين هذا الاختبار. وبالتحديد، يمكننا أن نجري اختبار النموذج في (6.73) عن طريق اختبار الافتراض المشترك في (6.77) أي $H_0: a_1 = a_2 = \dots = a_k = 0$ ، فإذا تم قبول H_0 فسوف نستنتج أن تباين u_t لا يعتمد على (X_{2t})، وحيث سنعتبر u_t يتسم بثبات التباين. أما إذا رفضنا H_0 فسوف نستنتج أن u_t تتسم باختلاف التباين ولذا سوف نواصل من خلال منهج التقدير السابق.*

ويمكننا بناء اختبار عينة كبيرة للافتراض H_0 من خلال إحداث تغير في المنهج المقترح من الملحق (ب) (B) للفصل الخامس. وبالتحديد (مرة أخرى) دع \hat{u}_t الخطأ العشوائي المقدر رقم t للنموذج (6.72) الذي يمكن الحصول عليه عن طريق تطبيق منهجنا المعتاد لذلك النموذج. دع ESS_u مجموع مربعات الخطأ الذي حصل عليه

* ينبغي علينا - نظرياً - قبل المضي في تصحيح مشكلة اختلاف التباين أن نحصل على عينة جديدة من المشاهدات. ولكن، في حالات عديدة، لا يمكننا القيام بذلك، ولذا فسوف نستمر في العمل بعينتنا الأولى.

من خلال تكوين انحدار \hat{u}_t^2 على عدد $(k+1)$ من المتغيرات في (6.77). أي الحد الثابت و X_{2t}, \dots, X_{kt} ، اجعل $\hat{\sigma}_t^2$ أيضاً، هو التقدير المناظر لتباين الخطأ العشوائي الذي حصل عليه بالطريقة العادية أي $(N-K-1) / ESS_u$ ، حيث إن N : حجم العينة. دع أخيراً ESS_R هو مجموع مربعات الخطأ الذي حصل عليه عن طريق تكوين انحدار \hat{u}_t^2 على الحد الثابت فقط. في هذه الحالة يكون $ESS_R = \sum_{t=1}^N (\hat{u}_t^2 - A)^2$ حيث إن $A = \sum_{t=1}^N \hat{u}_t^2 / N$. يمكننا حيثئذ أن نثبت أنه إذا كان $N = \infty$ فإن $(ESS_R - ESS_u) / \hat{u}_t^2$ سيكون متغير χ_k^2 بدرجات حرية قدرها k . وبطريقة مشابهة للمناقشة الموجودة في الملحق ب (B) من الفصل الخامس، فإنه يمكن إثبات أنه إذا كانت الفرضية H_0 غير صحيحة فسيكون الخطأ العشوائي متسماً باختلاف التباين من النوع الموجود في (6.77) كما أن ESS_R سيميل للكبر بالنسبة لـ ESS_u . لذلك فإن قيمة كبيرة للاحصائية χ^2 و $(ESS_R - ESS_u) / \hat{\sigma}_t^2$ ، ستؤدي إلى رفض H_0 . وإذا رمزنا للمتغير χ^2 بدرجات حرية عددها k بالرمز χ_k^2 . حيثئذ فإنه بافتراض مستوى معنوية للاختبار 0.5٪، سوف تعرف القيم الكبيرة للإحصائية χ^2 بأنها جميع القيم التي تزيد عن $\chi_{k, 0.95}^2$ ، حيث إن احتمال $[(\chi_k^2 \leq \chi_{k, 0.95}^2) = 0.95]$ ويمكن الحصول على قيمة $\chi_{k, 0.95}^2$ من أي جدول للمتغير χ^2 .

وبالطبع، وعند التطبيق، فإن عيتنا لن تكون ذات حجم لانهائي، لذلك، ينبغي اعتبار نتائج اختبارنا تقريبية. ونلاحظ (على سبيل معلومة جديرة بالاهتمام) أنه إذا كانت $N = \infty$ فإنه يمكن إثبات أن هذا الاختبار لـ χ^2 يكون معادلاً لاختبار F الذي يمكن بناءه عن طريق اتباع المنهج الموجود في الملحق ب B للفصل الخامس بالنسبة للمعادلة (6.77) وذلك بعد إحلال \hat{u}_t^2 محل u_t .

ونشير أخيراً إلى أنه، على الرغم من أننا قد أنشأنا اختباراً لاختلاف التباين بدلالة متغير مستقل واحد، فإنه من السهل تعميم هذا الاختبار ليشمل حالة كون عدد من المتغيرات المستقلة مصدراً لاختلاف التباين. افترض، على سبيل المثال، أن تباين الخطأ العشوائي في (6.72) يعتمد على كل من X_{1t} و X_{2t} ، أي:

$$\sigma'_{u_i} = g(X_{1i}, X_{2i}). \quad (6.79)$$

ومرة أخرى، على افتراض أن الدالة غير معلومة، فإن منهجنا أساسا سيكون مماثلا لما أوضحناه سابقا فيما عدا أن (6.77) سيتم إحلالها بمتعدد في كل من X_{1i} و X_{2i} . وللتوضيح، افترض أن $k=2$ ، حينئذ، سيتم إحلال المعادلة التالية محل المعادلة (6.7):

$$u_i^2 = a_0 + a_1 X_{2i} + a_2 X_{2i}^2 + b_1 X_{1i} + b_2 X_{1i}^2 + c_1 X_{1i} X_{2i} + \varepsilon_i. \quad (6.80)$$

سيرتبط اختبار ثبات التباين بفرضية العدم $H_0 : a_1 = a_2 = b_1 = b_2 = c_1 = 0$. وبالتحديد، دع ESS_u : مجموع مربعات الخطأ التي حصل عليها من انحدار \hat{u}_i^2 على الحد الثابت والمتغيرات $X_{1i}, X_{2i}, X_{1i}^2, X_{2i}^2, X_{1i} X_{2i}$. ولما كان لهذا الانحدار ست معاملات فإننا سنقدر σ_ε^2 على النحو: $\hat{\sigma}_\varepsilon^2 = ESS_u / (n-6)$. دع ESS_R (مرة أخرى): مجموع مربعات الخطأ الناتج من انحدار \hat{u}_i^2 على الحد الثابت فقط. يمكن في هذه الحال، إثبات أنه إذا كان حجم العينة لانهائيا فإن $\hat{\sigma}_\varepsilon^2 / (ESS_R - ESS_u)$ سيكون متغير χ^2 خمسة درجات حرية. لاحظ أن الذي يحدد ESS_u (في هذه الحالة) له خمس متغيرات مستقلة.* فإذا أخذنا مستوى المعنوية مساويا لـ 0.05 فسنفرض في هذه الحالة H_0 إذا تحققت $\hat{\sigma}_\varepsilon^2 / (ESS_R - ESS_u) > \chi_{5,0.95}^2$. ومرة أخرى، فإن نتائج هذا الاختبار تعد نتائج تقريبية فقط طالما أن حجم العينة، عادة، محدود من الناحية التطبيقية.

دعنا نذكر بعض الملاحظات الختامية لاختبارات التباين. إذا كانت درجة متعدد الحدود المقرب لنمط اختلاف التباين صغيرة ($k \geq 2$ مثلاً) فإن حدود التفاعل interaction terms التي تتضمن كلا المتغيرين المستقلين معا سوف تضمن في الانحدار. على سبيل المثال، يظهر الحد $X_{1i} X_{2i}$ في المعادلة (6.80) ولكن إذا كانت k كبيرة ($k \leq 3$ مثلاً) فإننا لن نهتم بجميع الحدود المتضمنة لأكثر من متغير مستقل واحد،

* عموماً، تكون درجات الحرية لمتغير χ^2 المناظرة لـ $(ESS_R - ESS_u) / \hat{\sigma}_\varepsilon^2$ مساوية لعدد المتغيرات المستقلة (غير

متضمنة الحد الثابت) المحددة لـ ESS_u .

وسبب ذلك هو أنه إذا كانت k كبيرة وكل حدود التفاعل الممكنة بين المتغيرات المستقلة متضمنة في النموذج. فإن عدد المتغيرات المستقلة سيكون كبيراً جداً، وقد يؤدي ذلك إلى ظهور مشاكل الارتباط الخطي المتعدد.

أما إذا كان حجم العينة لانهائياً، ولكن k ، درجة متعدد الحدود، محدودة (مهما كانت كبيرة) حيثئذ فإن كل التفاعلات الممكنة ينبغي، من حيث المبدأ، الاهتمام بها بهدف جعل اختبار اختلاف التباين جيداً بقدر الإمكان. وتعني عبارة «جيداً بقدر الإمكان» أنه عند مستوى معين من الخطأ من النوع الأول ينخفض الخطأ من النوع الثاني إلى حده الأدنى. ولكن، من الناحية العملية، يكون حجم العينة، عادة، محدوداً، ولذلك لا توجد -في هذه الحالة- نتائج محددة تشير إلى العدد الذي يجب أخذه من حدود التفاعل، وذلك لجعل الاختبار جيداً بقدر الإمكان. والتوجيه الوحيد الذي يمكن أن نقدمه في هذا المجال يكون التالي: دع p العدد الكلي للمتغيرات المستقلة في النموذج الذي يحدد مجموع مربعات الخطأ والمشار إليه بالرمز ESS ، حيثئذ، نقترح أن يكون عدد الحدود التي تختبر في متعدد الحدود حيث تتحقق غير المتساوية التالية $(n-p) \geq 25$ وهذا الاقتراح مبني على حدسنا فقط.

اختبار آخر لاختلاف التباين

اختبار جولدفيلد - وكوندات Goldfeld - Quandt

نناقش الآن اختباراً آخر لاختلاف التباين، وهذا الاختبار، في ظل تحقق شروط معينة يكون سهلاً ومشجعاً. افترض، بالتحديد، أننا نعتقد أن واحداً من المتغيرات المستقلة هو مصدر اختلاف التباين. افترض، أيضاً، أن العلاقة بين هذا المتغير وتباين الخطأ العشوائي مضطردة $monotonic$ ونعني بذلك أن تباين الخطأ العشوائي إما أن يتزايد باتساق مع قيمة المتغير المستقل، أو يتناقص باتساق مع قيمة المتغير المستقل. على سبيل المثال، للتوضيح اقترحنا في مناقشتنا السابقة في المعادلة (6.70) احتمال تزايد تباين حجم الأرباح مع زيادة قيمة أصول المنشأة.

وتمثل هذه إحدى حالات العلاقة المتزايدة باضطراب بين تباين الخطأ العشوائي وأحد المتغيرات المستقلة.

إذا كان الخطأ العشوائي مرتبطاً باضطراب مع أحد المتغيرات المستقلة وكان هذا الخطأ العشوائي موزعاً توزيعاً طبيعياً، فإنه يمكننا اختبار وجود اختلاف التباين باستخدام اختبار (جولدفيلد - كوندات)* (أو ج - ك)، هذا الاختبار له خصائص مشجعة. فبالمقارنة مع اختبارنا للعينات الكبيرة في المبحث السابق، فإن اختبار (ج - ك) هو اختبار للعينات الصغيرة. ولذا، ليست هناك ضرورة لاعتباره اختباراً تقريبياً للعينات الأقل من اللانهائية.

افترض وجود نموذج الانحدار الخطي التالي:

$$Y_i = b_0 + b_1 X_{1i} + \dots + b_h X_{hi} + \dots + b_k X_{ki} + u_i, \quad (6.81)$$

حيث X_h هو المتغير المستقل الذي نشك بأنه مصدر اختلاف التباين. افترض أن الخطأ العشوائي موزعاً توزيعاً طبيعياً، وإذا وجد اختلاف التباين، فسببه ارتباط تباين الخطأ العشوائي باضطراب مع X_h . بافتراض تحقق هذه الافتراضات، يتبع منهج اختبار (ج - ك) الخطوات التالية:

١- رتب جميع المشاهدات وفقاً لقيم X_h . فإذا كانت العلاقة المشكوك فيها بين X_h وتباين الخطأ العشوائي علاقة مضطربة وموجبة، حينئذ ينبغي أن تكون أول مشاهدة أعيد ترتيبها تناظر أصغر قيمة لـ X_h . وتناظر المشاهدات التالية ثاني أصغر مشاهدة لـ X_h وهلم جرا. وعلى العكس إذا كانت العلاقة المشكوك فيها بين X_h وتباين الخطأ العشوائي علاقة مضطربة وسالبة، ينبغي أن تكون أول مشاهدة أعيد ترتيبها تناظر أكبر قيمة لـ X_h ، وتناظر المشاهدات الثانية ثاني أكبر مشاهدة وهلم جرا. إذا رتبنا المشاهدات بهذه الطريقة، ووجد اختلاف التباين فإن تباين الأخطاء العشوائية i, j سوف يكون حيث $\text{var}(u_i) < \text{var}(u_j)$ إذا كانت $i < j$ على سبيل

* يرجع إلى:

التوضيح لإعادة الترتيب، افترض النموذج التالي:

$$Y_t = a_0 + a_1 X_{1t} + \dots + a_2 X_{2t} + u_t, \quad t=1,2,\dots,6 \quad (6.82)$$

بافتراض أن $\text{var}(u_t)$ يزداد بزيادة قيمة X_{1t} ، حينئذ إذا كانت العينة الأولية هي:

	Y	X_1	X_2
t=1	10	2	15
t=2	12	1	27
t=3	-1	10	0
t=4	5	9	5
t=5	3	27	1
t=6	0	5	10

(6.83)

فإن العينة المعاد ترتيبها هي:

Y	X_1	X_2
12	1	27
10	2	15
0	5	10
5	9	5
-1	10	0
3	27	1

(6.84)

أما إذا فرض أن u_t تباين u_t يتناقص مع زيادة قيمة X_{1t} ، فإن العينة المعاد ترتيبها

ستكون على النحو التالي:

Y	X_1	X_2
3	27	1
-1	10	0
5	9	5
0	5	10
10	2	15
12	1	27

(6.85)

٢- احذف عدد d من المشاهدات من وسط العينة المعاد ترتيبها. وعلى الرغم من أن الرقم d هو رقم تحكيمي، إلا أن أخذ $d = n/3$ ، بالتقريب، يعد قاعدة حسابية

معقولة حيث إن n هو حجم العينة الأولية، يضاف إلى ذلك أن d ينبغي أن تختار حيث تكون $(n-d)$ عددا صحيحا زوجيا، على سبيل المثال، إذا كانت $n=61$ تكون $d=21$ لذلك يكون $n-d=40$ وهو رقم زوجي، فإذا تم ذلك فإن العينة الأولية ستجزأ إلى عيتين فرعيتين يحتوي كل منهما على $(n-d)/2$ من المشاهدات.

وللتوضيح، إذا كانت المعادلة (6.83) هي العينة الأولية والمعادلة (6.84) هي العينة المعاد ترتيبها، حيث، تكون $d=6/3=2$ ، ولذلك، سيتم إسقاط المشاهدين الوسيطين في المعادلة (6.84). وتكون العينة الفرعية الأولى هي أول $(n-d)/2 = (6-2)/2 = 2$ مشاهدين.

Y	X ₁	X ₂
12	1	27
10	2	15

(6.86)

وستكون العينة الفرعية الثانية:

Y	X ₁	X ₂
-1	10	0
3	27	1

(6.87)

٣- قدر معادلات انحدار منفصلة لكل من العيتين الفرعيتين.

٤- احسب مجموع مربعات الخطأ (ESS) لكل واحدة من معادلات الانحدار. ESS_1 هو ESS للعينة الفرعية الأولى و ESS_2 ليكون ESS للعينة الفرعية الثانية. لاحظ أنه، إذا وجد اختلاف في التباين، تكون الأخطاء العشوائية للعينة الفرعية الأولى ذات تباين أقل من تباين العينة الفرعية الثانية.

٥- لتبسيط الرموز، دع p عدد معاملات الانحدار (على سبيل المثال $(P=K+1)$)

في المعادلة (6.81) حيث يمكن (ESS_2/ESS_1) إثبات هي موزعة بالضبط باعتبارها F بدرجات حرية عددها $(n-d-2p)/2$ في كل من البسط والمقام. وبإجراء الخطوات السابقة فإننا سنرفض فرضية العدم H_0 ، ثبات التباين للأخطاء العشوائية عند مستوى المعنوية المختار، إذا كانت (ESS_2/ESS_1) أكبر من القيمة الحرجة لتوزيع F كما توجد في جدول F .

وعلى سبيل التوضيح، دع $e = (n-d-2p)/2$ ، ودعنا نرسم إلى المتغير F الذي له درجات حرية عددها e في كل من البسط والمقام بالرمز $F_{e,e}$. دع $F_{e,e}^{0.95}$ هي قيمة F من الجدول حيث يكون احتمال $F_{e,e} \leq F_{e,e}^{0.95} = 0.95$. حيث إذا أخذنا مرة أخرى، مستوى المعنوية ليكون (0.95) فإننا سنرفض H_0 إذا كانت $ESS_2/ESS_1 > F_{e,e}^{0.95}$. وهكذا، فإن اختبار (ج - ك) اختبار بسيط ومباشر. إضافة إلى ذلك، فإن منطق هذا الاختبار واضح. فبديها، إذا كان تباين الخطأ العشوائي مرتبطاً، حقيقة، بالمتغير المستقل (المشكوك فيه)، فإنه ينبغي علينا توقع أن تكون مربعات الأخطاء العشوائية المقدرة أكبر في العينة الفرعية الثانية منه في الأولى، لذلك، فإن القيم الأكبر لـ (ESS_2/ESS_1) تؤدي إلى رفض فرضية العدم، أما إذا كانت الأخطاء العشوائية المقدرة لها تقريبا الحجم في العيتين الفرعيتين، حيث، تقرب النسبة (ESS_2/ESS_1) من الواحد الصحيح، وحينها سنقبل فرضية العدم (أي ثبات التباين للأخطاء العشوائية) لأن القيمة الحرجة لاختبار F تكون أكبر من الواحد الصحيح عند مستوى المعنوية 0.05 أو 0.01 ولدرجات حرية متساوية لكل من البسط والمقام. هناك نقطتان ختاميتان ينبغي ملاحظتهما. أولاها: يكون اختبار (ج - ك) صحيحا حتى إذا لم تحذف مشاهدات من وسط العينة المعاد ترتيبها. ولكن، أوضحت التجارب أن إلغاء بعض المشاهدات يحسن الاختبار لأنه يؤدي إلى تخفيض حجم الخطأ من النوع الثاني. والنقطة المهمة هي أن المشاهدات المحذوفة تجعل القسم الأول والقسم الثاني من العينة المعاد ترتيبها أكثر اختلافاً، وهكذا يصبح من السهل اكتشاف الاختلافات في التباين للأخطاء العشوائية في العيتين الفرعيتين. والنقطة الثانية هي أن مناقشتنا السابقة قد افترضت، ضمناً، أن $(n-d)/2 > p$ أما إذا كانت $(n-d)/2 < p$ ، فإن الاختبار لا يمكن تنفيذه لأنه لن تكون هناك مشاهدات كافية في كل قسم من العينة لتقدير معلمات نموذج الانحدار، ومن ثم، لتحديد ESS_1 و ESS_2 إذا كان $(n-d)/2 = p$ يكون عدد المشاهدات في كل قسم من العينة مساوياً لعدد المعلمات التي تقدر. في هذه الحال، يمكن إثبات أنه إذا لم يكن هناك تعدد خطي تام بين المتغيرات المستقلة فإن كلا من ESS_1 و ESS_2 سيساوي الصفر، وهكذا

لا يمكن، مرة أخرى، إجراء الاختبار. وقد افترضنا أنه إذا كانت d تعادل $n/3$ وتبين حينئذ أن $p \leq (n-d)/2$ ، حينئذ لا ينبغي إسقاط أي مشاهدة من منتصف العينة المعاد ترتيبها (أي بجعل d مساوية للصفر). فإذا تم ذلك وكانت n عددا زوجيا فسيكون هناك $n/2$ من المشاهدات في كل من العيتين المعاد ترتيب كل منهما، أما إذا كانت $n/2 > p$ فإن الاختبار ينبغي أن يجري كما شرحنا من قبل. وفي هذه الحال، ينبغي رفض H_0 إذا كان مستوى المعنوية 0.05 و $ESS_2/ESS_1 > F_{(n-2p)/2, (n-2p)}^{0.95}$. فإذا كانت n رقما فرديا نقترح جعل $n_1 = (n+1)/2$ في العينة الفرعية الأولى و $n_2 = n - n_1$ للعينة الفرعية الثانية. وهنا ينبغي أن يجرى الاختبار كما في السابق باستثناء أننا سنرفض H_0 إذا كانت: $ESS_2/ESS_1 > F_{((n_1-p)/2, (n_2-p)/2)}^{0.95}$ ، وأخيرا، إذا كانت $n/2 \leq p$ فإنه لا ينبغي استخدام اختبار (ج - ك) للكشف عن اختلاف التباين.

بعض التعليقات حول اختباري اختلاف التباين

افترضنا، حتى الآن، اختبارين لاختلاف التباين. ولكن هناك اختبارات أخرى، وسبب ذلك هو أنه لا يوجد حتى الآن نموذج واحد يصلح لكل الظروف المحتملة، ففي ظروف معينة، يكون من الملائم استخدام اختبار معين، بينما في ظروف أخرى، يكون من الملائم استخدام اختبار آخر. وينبغي على الباحث أن يكون قادرا على استخدام هذين الاختبارين اللذين افترضناهما للتطبيق في معظم الحالات.

ولمعرفة القضايا المتضمنة في هذه الاختبارات، تذكر أنه، مع تحقق الافتراضات السابقة، فإن اختبار (ج - ك) مفيد لكونه اختبار عينة صغيرة، فليس مبنا على تقريب للحالة المتضمنة عينة لانتهائية. لذلك، فإن نتائجه دقيقة، وليس تقريبية. على سبيل المثال، فإن استخدام اختبار (ج - ك) عند مستوى معنوية 0.05 له في الحقيقة، خطأ من النوع الأول مساو لـ 0.05 . إضافة إلى ذلك فإن التجارب قد أظهرت - مع تحقيق الافتراضات اللازمة لاستخدام الاختبار - أن الخطأ من النوع الثاني لاختبار (ج - ك) صغير بشكل مقبول. لذلك، إذا تحققت الافتراضات سالفة الذكر، فإن اختبار (ج - ك) يعد اختبارا جيدا للاستخدام.

ولكن، قد لا تتحقق الافتراضات سالفة الذكر. فقد تكون الأخطاء العشوائية، مثلا، موزعة توزيعا غير طبيعي، حيثئذ ينبغي أخذ نتائج اختبار (ج - ك) على أنها نتائج تقريبية. وبديها، تقترب النتائج من الحقيقة كلما اقترب توزيع الأخطاء العشوائية من التوزيع الطبيعي. وفي الحقيقة، يرغب الاقتصاديون عند التطبيق افتراض أن الأخطاء العشوائية موزعة توزيعا طبيعيا، ولذا، لا يكون هذا الافتراض محل اهتمام كبير من جانبهم، وبالطبع، فنحن نذكر، فقط، وجهة النظر هذه دون أن نتبناها. والافتراض الأكثر خطورة للمنهج ذاته هو افتراض أنه إذا وجد اختلاف التباين فهو بسبب ارتباط تباين الخطأ العشوائي باضطراب بواحد من المتغيرات المستقلة. هذا الافتراض يمكن الباحث من إعادة ترتيب العينة بطريقة لن يتناقص معها تباين الخطأ العشوائي وهذا هو حجر الزاوية للاختبار. ومن الواضح، أنه إذا كان تباين الخطأ العشوائي مرتبطا بأكثر من متغير مستقل واحد، أو كان مرتبطا بمتغير مستقل واحد ولكن بطريقة أخرى غير مضطربة، حيثئذ، لا يستطيع الباحث عموما أن يعيد ترتيب العينة كما شرحنا من قبل، ومن ثم، لن يتمكن من إجراء هذا الاختبار. يمكن للباحث في مثل هذه الحالات أن يستخدم الاختبار الذي شرحناه في المبحث السابق. ذلك الاختبار ليس محددا بالحالات التي يكون فيها واحد من المتغيرات المستقلة، فقط، هو السبب في ظهور مشكلة اختلاف التباين، كما أنه لا يتطلب أن تأخذ العلاقة بين المتغيرات المستقلة المتسببة في اختلاف التباين وتباين الأخطاء العشوائية الصورة المضطربة. وفي الحقيقة، يمكن أن يتناول هذا الاختبار أنماطا معقدة جدا من اختلاف التباين، وذلك من خلال جعل درجة متعدد الحدود المقرب، مثلا، في المعادلة (6.73) كبيرة. نلاحظ، أخيرا أن ذلك الاختبار ليس مبني على افتراض كون الأخطاء العشوائية موزعة توزيعا طبيعيا.

ولكن هذا الاختبار له حدوده، أيضا، فهو اختبار للعينات الكبيرة، ولما كانت العينات التي يستخدمها الباحث محدودة، فإنه لن يكون متأكدا من خصائص الاختبار الذي يستخدمه - على سبيل المثال حجم الخطأ من النوع الأول! وأيضاً، يكون حجم الخطأ من النوع الثاني - في هذا الاختبار كبيرا إذا كان حجم العينة صغيرا.

اختلاف التباين: نتيجة التجميع

سوف ننهي معالجتنا لاختلاف التباين بحال واقعية ظهرت في مجال تقدير الطلب على القمح في الولايات المتحدة الأمريكية*. ويختلف مصدر اختلاف التباين في هذه الحالة عن المصادر التي عالجناها من قبل. وبشكل خاص تبدأ الدراسة بنوع معتاد من دوال الطلب:

$$Q_t = b_0 + b_1 P_t^w + b_2 P_t^s + b_3 Y_t + b_4 D_t + b_5 S_{1t} + b_6 S_{2t} + b_7 S_{4t} + u_t \quad (6.88)$$

في هذه المعادلة، يعتمد الطلب على القمح في الفترة t (أو Q_t) على الأسعار الجارية للقمح (P_t^w) وأسعار الحبوب الأخرى (P_t^s)، ومتوسط دخل الفرد (Y_t)، وأخيراً، على أربعة متغيرات صورية. يأخذ المتغير الصوري الأول منها (D_t) (والذي يأخذ في الحسبان تكلفة شهادات تسويق معينة ينبغي أن يشتريها المتعاملون في الصناعات الغذائية المحلية خلال جزء من الفترة موضع الاهتمام، ويأخذ قيمة الواحد الصحيح خلال تلك الفترات التي تكون فيها الشهادات فعالة، وقيمة الصفر في الأوقات الأخرى). أما باقي المتغيرات الصورية S_1, S_2, S_4 ، فهي متغيرات صورية موسمية لفصول: الأول والثاني والرابع من السنة الميلادية على التوالي.

ومصدر المشكلة في هذه الحالة هو طبيعة البيانات، حيث تقدم وزارة الزراعة الأمريكية أرقاماً كل فترة زمنية عن الكميات والأسعار للقمح والحبوب الأخرى. وتتوافر البيانات عن هذه المتغيرات على أساس ربع سنوي بدءاً من الربع الثاني من سنة ١٩٦٤م، ولكن البيانات المتاحة قبل هذا التاريخ والمرتبطة بالطلب على القمح تتوافر، فقط، على أساس نصف سنوي.

على سبيل المثال، فإن المشاهدة المتاحة عن الطلب قبل الربع الثاني عام ١٩٦٤م مرتبطة بالنصف الأول (أول فصلين ربع سنويين) لسنة ١٩٦٤م، والمشاهدة المتاحة قبل هذه مرتبطة بالنصف الثاني (آخر فصلين ربع سنويين) لسنة ١٩٦٣م

* اشترك في هذه الدراسة واحد من المؤلفين، انظر:

David Bradford and Harry H. Kelejian, A Quarterly Demand Model for Wheat (Unpublished Manuscript, 1976).

وهلم جرا. هناك مشكلة واضحة إذا أردنا استخدام البيانات قبل الفصل الثاني لعام ١٩٦٤م وبعده لتقدير دالة طلب موحدة. كيف يمكننا استخدام البيانات نصف السنوية وربيع السنوية لتقدير معادلة انحدار واحدة؟ باستخدام نموذج أكثر عمومية، افترض أن نموذج الانحدار الموضح للقيم ربع السنوية للمتغير التابع Y_t هو:

$$Y_t = a_0 + a_1 X_{1t} + \dots + a_n X_{nt} + u_t, \quad (6.89)$$

حيث يشير الرمز السفلي t إلى الفصول ربع السنوية من السنة الميلادية، وحيث إن الخطأ العشوائي u_t يحقق الافتراضات المعتادة كامة. افترض - بالتحديد - أن u_t مستقل عن جميع القيم الماضية والحالية والمستقبلية للمتغيرات المستقلة وغير مرتبط ذاتيا، وله قيمة متوقعة مساوية للصفر $E(u_t) = 0$ ، وأخيرا، تباين ثابت $E(u_t^2) = \sigma_u^2$.

افترض أن المشاهدات ربع السنوية عن المتغير التابع متاحة، فقط، للفترات $(t = T, T+1, T+2, \dots, T+N)$. أما بالنسبة للفترات السابقة عن الفترة T ، فإنه توجد لدينا، فقط، مشاهدات نصف سنوية ليست متداخلة $(Y_{T-2} + Y_{T-1})$ ، $(Y_{T-4} + Y_{T-3})$ ، ...، $(Y_{T-\phi} + Y_{T-\phi-1})$ ، حيث إن ϕ هو رقم صحيح فردي يشير إلى عدد المشاهدات نصف السنوية المتاحة - وسنناقش هذا أدناه. في المثال السابق، ترتبط $(Y_{T-2} + Y_{T-1})$ بالنصف الأول من عام ١٩٦٤م، بينما ترتبط $(Y_{T-4} + Y_{T-3})$ بالنصف الثاني من عام ١٩٦٣م وهلم جرا. وبالعودة إلى نموذجنا (6.89)، نفترض أن المشاهدات ربع السنوية تكون متاحة لكل متغير مستقل ولجميع الفترات موضع الاهتمام (في المثال السابق - مثلا - قبل الفصل الثالث لسنة ١٩٦٤م وبعده).

إذا كان (6.89) هو نموذج الانحدار الذي يفسر المتغير ربع السنوي Y_t ، فإنه يترتب على ذلك أن نموذج الانحدار للمتغير نصف السنوي $(Y_{T-j} + Y_{T-j-1})$ يكون:

$$(Y_{T-j} + Y_{T-j-1}) = 2a_0 + a_1(X_{1,T-j} + X_{1,T-j-1}) + \dots + a_n(X_{n,T-j} + X_{n,T-j-1}) + (u_{T-j} + u_{T-j-1}), \quad j=1, 3, 5, \dots, \phi, \quad (6.90)$$

حيث ϕ هو رقم فردي صحيح يحدد قيمته (كما سنوضح فيما بعد) عدد المشاهدات نصف السنوية المتاحة. وقد امكن الحصول على النموذج (6.90) من خلال جمع الجانب الأيمن من (6.89) للفصول ربع السنوية التي تناظر المتغير التابع، أي $T-j$ و $T-j-1$.

وفي ظل تحقق افتراضاتنا، يكون للخطأ العشوائي $(u_{T-j} + u_{T-j-1})$ في النموذج نصف السنوي (6.90) قيمة متوقعة صفرية، وغير مرتبط ذاتياً، طالما أن المكونات ربع السنوية ليست متداخلة، ومستقلة عن القيم الماضية والجارية والمستقبلية كافة، ولها تباين ثابت وعلى وجه التحديد:

$$E(u_{T-j} + u_{T-j-1})^2 = 2\sigma_u^2. \quad (6.91)$$

لذلك، يحقق نموذجنا (6.90) المرتبط بالملاحظات نصف السنوية جميع الافتراضات المعتادة، ولكن النموذج ربع السنوي (6.90) يحقق، أيضاً، جميع الافتراضات المعتادة، ويحتوي على المعلومات غير المعلومة نفسها. سوف نبين الآن أن هذين النموذجين يمكن أن يدمجا معا في نموذج واحد يحقق افتراضاتنا المعتادة كافة، ويمكن استخدام كل من البيانات ربع السنوية ونصف السنوية لتقدير معالمته. لاحظ أن مشاهداتنا المتاحة عن المتغير التابع يمكن ترتيبها زمنياً من الأقدم إلى الأحدث كمايلي:

$$(Y_{T-\phi} + Y_{T-\phi+1}), (X_{T-\phi+1}), \dots, (X_{T-5} + X_{T-6}), (Y_{T-3} + Y_{T-4}), \\ (Y_{T-1} + Y_{T-2}), Y_T, Y_{T+1}, \dots, Y_{T-N}. \quad (6.92)$$

ولاحظ أنه إذا كانت $\phi = 5$ فسوف يكون لدينا ثلاث مشاهدات نصف سنوية هي $(Y_{T-6} + Y_{T-5})$ ، $(Y_{T-4} + Y_{T-3})$ ، $(Y_{T-2} + Y_{T-1})$. لاحظ، أيضاً، أن $(5+1)/2 = 3$. وعلى سبيل مثال آخر ينبغي أن يكون واضحاً أنه إذا كانت $\phi = 3$ فسيكون لدينا مشاهدتان نصف سنويتين. لاحظ، مرة أخرى، أن $(3+1)/2 = 2$. وعموماً توضح هذه الأمثلة أن المشاهدات نصف السنوية لأي رقم صحيح فردي يكون $(\phi + 1)/2$. لذلك يكون العدد الإجمالي للمشاهدات الموجودة في المعادلة (6.92) هو $\{[(\phi + 1)/2] + [N + 1]\}$.

دعنا الآن نرمز للملاحظات $\{[(\phi + 1)/2] + [N + 1]\}$ في المعادلة (6.92) بالرمز y_t حيث إن $\{[(\phi + 1) / 1] + [N + 1]\}$ وأن $t=1,2,\dots$. أي أن y_t تشير إلى أقدم مشاهدة في (6.92). أي $(Y_{T-\phi} + Y_{T-\phi-1})$ ، y_2 تشير إلى المشاهدة التالية وهلم جرا. وبطريقة مشابهة نجد أن عدد المشاهدات نصف السنوية لكل متغير مستقل في المعادلة (6.90)* يبلغ $(\phi + 1)/2$. إضافة إلى ذلك، فقد افترضنا أن المشاهدات عن كل متغير مستقل تكون متاحة لكل فترة من الفترات التي تتاح فيها بيانات ربع سنوية للمتغير التابع أي للفترات $T+N,\dots,T+1,T$. دع X_{jt} حيث $\{[(\phi + 1)/2] + [N + 1]\}$ نصف السنوية وربع السنوية $t=1,\dots$ يرمز إلى المشاهدات $\{[(\phi + 1)/2] + [N + 1]\}$ نصف السنوية وربع السنوية للمتغير المستقل X_{jt} التي رتبب زمنيا ترتيبا يشبه الطرق الموجودة في المعادلة (6.92). حيثئذ، تتضمن نتائجنا في المعادلة (6.89) والمعادلة (6.90) أن y_t مرتبطة بـ X_{2t} ، X_{1t} ، X_{nt} ،... على النحو التالي:

$$y_t = a_0 X_{0t} + a_1 X_{1t} + \dots + a_n X_{nt} + v_t, \quad t=1,2,\dots, \left(\frac{\phi+1}{2}\right) + (N+1), \quad (6.93)$$

حيث $x_{0t} = 2$ إذا كانت t تناظر مشاهدة نصف سنوية، أي أن $t \geq (\phi + 1)/2$ و $x_{0t} = 1$ فيما عدا ذلك. وحيث إن u_t هو الخطأ العشوائي الذي يحقق جميع الافتراضات النمطية فيما عدا أنه يتسم باختلاف التباين. وبالتحديد فإن $E(v_t^2) = 2\sigma_u^2$ إذا كانت $t \geq (\phi + 1)/2$ ، وأن $E(v_t^2) = 2\sigma_u^2$ في غير ذلك من الحالات. يمكن تحويل النموذج (6.93) إلى نموذج يحقق افتراضاتنا المعتادة كافة. بالتحديد، دع $d = \sqrt{2}$ إذا كانت $t \geq (\phi + 1)/2$ و $d_t = 1$ إذا كانت t غير ذلك. حيثئذ، يمكننا، باتباع المنهج الذي سبق توضيحه في المناقشات السابقة. أن نلغي مشكلة اختلاف التباين من خلال قسمة (9.93) على d_t وينتج عن ذلك النموذج:

* بحكم أننا افترضنا توافر المشاهدات ربع السنوية للمتغيرات المستقلة، فإنه من السهل تكوين المشاهدات نصف السنوية لتلك المتغيرات.

$$\left(\frac{y_t}{d_t}\right) = a_0 \left(\frac{x_{0t}}{d_t}\right) + a_1 \left(\frac{x_{1t}}{d_t}\right) + \dots + a_n \left(\frac{x_{nt}}{d_t}\right) + w_t, \quad (6.94)$$

$$t = 1, 2, \dots, \left(\frac{\varphi+1}{2}\right) + (N+1),$$

حيث إن $v_t / d_t = w_t$ ومن الواضح أن $E(w_t^2) = \sigma_v^2$ لجميع $t \in \{1, 2, \dots, [(\varphi+1/2)] + [N+1]\}$ ،

والآن يستوفي النموذج (6.94) افتراضاتنا المعتادة كافة، ويرتبط بجميع مشاهداتنا المتاحة نصف السنوية وربع السنوية عن المتغير التابع. ويمكن تقدير هذا النموذج بوساطة منهجنا المعتاد في التقدير. وبالتحديد، نجد أن المعادلات الطبيعية يمكن اشتقاقها بوساطة الشروط:

$$\sum \hat{w}_t \left(\frac{x_{0t}}{d_t}\right) = 0, \dots, \sum \hat{w}_t \left(\frac{x_{nt}}{d_t}\right) = 0, \quad (6.95)$$

حيث يتم كل تجميع على مدى $t \in \{1, 2, \dots, [(\varphi+1/2)] + [N+1]\}$ ، وحيث تكون $\hat{w}_t = (y_t / d_t) - \hat{a}_0 (x_{0t} / d_t) - \dots - \hat{a}_n (x_{nt} / d_t)$ ، حيث $\hat{a}_n, \dots, \hat{a}_0$ مقدرات a_n, \dots, a_0 على الترتيب.

نلاحظ أنه عندما يحصل على المقدرات $\hat{a}_n, \dots, \hat{a}_0$ فإنه يمكن تفسير y_t بوساطة النموذج على النحو التالي:

$$\left(\frac{\hat{y}_t}{d_t}\right) = \hat{a}_0 \left(\frac{x_{0t}}{d_t}\right) + \dots + \hat{a}_n \left(\frac{x_{nt}}{d_t}\right), \quad (6.96)$$

أو بإلغاء d_t على النحو:

$$\hat{y}_t = \hat{a}_0 x_{0t} + \dots + \hat{a}_n x_{nt} \quad (6.97)$$

ويعني هذا أن أحدث القيم ربع السنوية تفسر على النحو:

$$\hat{Y}_t = \hat{a}_0 + \hat{a}_1 X_{1t} + \dots + \hat{a}_n X_{nt}, \quad t = T, T+1, \dots, T+N; \quad (6.98)$$

بينما تفسر القيم نصف السنوية على النحو:

$$(\hat{Y}_{T-j} + \hat{Y}_{T-j-1}) = 2\hat{a}_0 + \hat{a}_1 (X_{1,T-j} + X_{1,T-j-1}) \\ + \dots + \hat{a}_n (X_{n,T-j} + X_{n,T-j-1}), \quad j=1,3,5,\dots,\phi.$$

(٤-٦) مشاكل في اختيار المتغيرات

افترضنا، حتى الآن، أن متغيرات نموذج الانحدار تعطي لنا بطريقة أو بأخرى، وأن مشكلتنا الوحيدة تنحصر في تقدير النموذج واختبار الفرضيات وعلاج الارتباط الذاتي وهلم جرا. ولكن، عند التطبيق، نجد أنه ينبغي علينا اختيار المتغيرات التي يشتمل عليها النموذج. ويعتمد الباحث في تحديده للمتغير التابع على إحدى النظريات، ثم يحاول، بعد ذلك، تحديده المتغيرات المستقلة التي تفسر النظرية تفسيراً أفضل. وعند القيام بذلك يمكن أن يقع الباحث في نوعين من الأخطاء. أولاً: قد يفشل الباحث في تضمين أحد المتغيرات المستقلة مهمة في نمودجه، بمعنى أنه قد يغفل أحد العوامل المهمة المحددة للمتغير التابع. ثانياً قد يفترض الباحث أن عاملاً معيناً يكون مهماً في تحديد المتغير التابع في الوقت الذي لا يكون فيه ذلك العامل مهماً في الحقيقة. فإذا ما حدث ذلك فسيشتمل نمودجه على متغير غير ضروري. سنهتم، في هذا البحث، بمناقشة نتائج الوقوع في كل واحد من هذه الأخطاء.

متغير محذوف

دعنا نهتم أولاً بحالة حذف المتغيرات المستقلة من العلاقة المفترضة، افترض، على سبيل المثال، أن العلاقة الحقيقية (غير المعلومة لنا) هي:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t, \quad (6.99)$$

حيث u_t الخطأ العشوائي الذي يستوفي جميع افتراضاتنا المعتادة، ولكننا أغفلنا X_2 واعتبرنا أن المعادلة تأخذ الشكل التالي:

$$Y_t = b_0 + b_1 X_{1t} + r_t, \quad (6.100)$$

حيث جعلنا r_t خطأنا العشوائي. سنبين الآن أنه إذا أردنا تقدير المعادلة (6.100) بطريقتنا العادية، فإن المقدرات الناتجة لكل من b_0 و b_1 سوف تكون، عموماً، متحيزة وغير متسقة. والسبب هو أن الفشل في إدخال X_2 في النموذج يؤدي إلى خرق لافتراضاتنا الأساسية المرتبطة بالخطأ العشوائي، وهو في هذه الحالة، r_t . وبالتحديد وبمقارنة المعادلة (6.99) مع المعادلة (6.100) سنرى أن الخطأ العشوائي

$$\text{- في النموذج الذي يجري تقديره - يعتمد جزئياً على } X_{2t}. \\ r_t = b_2 X_{2t} + u_t, \quad (6.101)$$

وبأخذ القيم المتوقعة نجد:

$$E(r_t) = E(b_2 X_{2t}) + E(u_t) = b_2 \mu_2 + 0 = b_2 \mu_2, \quad (6.102)$$

حيث μ_2 هي القيمة المتوسطة لـ X_{2t} . ومن الواضح أن القيمة المتوقعة للخطأ r_t لن تكون، عموماً، مساوية للصفر، فيما عدا الحالة التي تكون فيها $b_2 = 0$ ، والتي تعني أن Y_t لا تعتمد على X_{2t} .^{*} وهكذا، نرى أنه إذا كانت Y دالة في X_2 في معادلة الانحدار، فإن الخطأ العشوائي في تلك المعادلة لن يكون، له عموماً، قيمة متوسطة صفرية. يضاف إلى ذلك أنه ينبغي أن يكون واضحاً ارتباط r_t بـ X_{1t} إذا كانت X_{2t} مرتبطة بـ X_{1t} ، ونتيجة لذلك، فإن التباين $\text{cov}(r_t, X_{1t})$ لن يكون، عموماً صفراً. ويتبع عن ذلك، في الأقل، بديهياً، في ظل مثل هذه الشروط أن مقدراتنا لـ b_0 و b_1 ستكون متحيزة وغير متسقة. فمثلاً، إذا كانت $E(r_t) \neq 0$ ولكننا حصلنا على معادلتنا الطبيعية الأولى عن طريق وضع $\hat{r} = 0$ ، فإن منهجنا في التقدير لن يتسم بالاتساق.

وقد يكون من المفيد أن نحدد طبيعة هذا التحيز عند مستوى تحليلي أعمق. افترض أننا نرغب في تقدير دالة استهلاك وأن العلاقة الحقيقية تأخذ الشكل:

$$C_t = b_0 + b_1 Y_{dt} + b_2 A_t + u_t, \quad (6.103)$$

^{*} والحالة الأخرى المناظرة هي عندما يكون $E(r_t) = 0$ ، وهي حالة عرضية وفريدة لأنها تعني أن $\mu_2 = 0$.

حيث:

$$\begin{aligned} C_t & \text{ الانفاق الاستهلاكي خلال الفترة الزمنية } t, \\ Y_{dt} & \text{ الدخل المتاح خلال الفترة الزمنية } t, \\ A_t & \text{ حجم الأصول السائلة في الفترة } t, \\ u_t & \text{ الخطأ العشوائي} \end{aligned}$$

نتوقع أن يكون لـ A_t تأثير على C_t ، ولذلك، فإن $b_2 > 0$. افترض، أيضا، أنه توجد علاقة ارتباط موجب بين A_t و Y_{dt} ، بمعنى أنه كلما تزايد الدخل المتاح تزايد أيضا، قيمة الأصول السائلة (والعكس صحيح). ولكن افترض أن المعادلة التي نقدرها فعلا لا تحتوي على A_t :

$$C_t = b_0 + b_1 Y_{dt} + r_t. \quad (6.104)$$

لا تكون القيمة المتوقعة للخطأ العشوائي في هذه الحالة (كما في التحليل السابق) للمعادلة التي اخترنا تقديرها مساوية للصفر، عموماً:

$$E(r_t) = E(b_2 A_t) + E(u_t) = b_2 \mu_A \neq 0, \quad (6.105)$$

حيث إن μ_A هي القيمة المتوسطة لـ A_t . وبالمثل، يمكن أن نتوقع أن يرتبط r_t بـ Y_{dt} . وسوف نحصل، لذلك، على مقدرات متحيزة لـ b_0 و b_1 (حيث يشير الأخير إلى م ح س). أكثر من ذلك، فإنه، بسبب التأثير الموجب لـ A_t على C_t ($b_2 > 0$) والارتباط الموجب بين A_t و Y_{dt} ، فسوف نحصل على تقدير أعلى مما يجب لـ م ح س، بمعنى أن $E(\hat{b}_1) > b_1$. وسبب ذلك، بديهياً، هو أن Y_{dt} في المعادلة (6.104) تعمل باعتبارها متغيراً ينوب عن نفسه وعن A_t . أي أنه ينسب إليه التأثير الموجب لكل من Y_{dt} و A_t على C_t . والنقطة المهمة هنا هي أنه، عندما يرتفع Y_{dt} فإن A_t ترتفع كذلك ولكن غياب A_t من المعادلة المقدرة ينسب التأثير الموجب لـ A_t على C_t إلى Y_{dt} ، ونتيجة لذلك، فإن التأثير المنسوب لـ Y_{dt} على C_t ارتباطاً سالباً ولكنها مازال مرتبطة بـ Y_{dt} ارتباطاً موجباً، حيثُ، فإن مقدرنا لـ b_1 في غياب A_t من معادلة الانحدار سيكون متحيزاً لأسفل، وسينعكس الأثر السالب لـ A_t على C_t ، بعض الشيء، في \hat{b}_1 . نلاحظ، أخيراً، أن اتجاه التحيز في هذه

الحالات سيكون (عمومًا) معكوسا إذا كانت Y_{dt} و A_{dt} مرتبطتين ببعضهما ارتباطا سالبًا. والنقطة المهمة في كل هذا ليست هي تحديد اتجاه التحيز بقدر ماهي إدراك أنه إذا حذف أحد المتغيرات المهمة من معادلتنا فإن مقدراتنا سوف تكون متحيزة وغير متسقة. وحقيقة فإنه لا يمكن، عادة، تحديد اتجاه التحيز في معظم نماذج الانحدار التي تشتمل على عدد كبير من المتغيرات المستقلة.

متغيرات أكثر من اللازم

دعنا الآن نهتم بحالة وجود عدد كبير من المتغيرات المستقلة في النموذج. افترض، على سبيل المثال، أن العلاقة الحقيقية تأخذ الشكل:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t, \quad (6.106)$$

حيث نفترض أن الخطأ العشوائي مستقل عن قيم جميع المتغيرات المستقلة، وأنه يحقق الافتراضات المعتادة الأخرى كافة. تبين هذه المعادلة أنه ليس من الضروري أن ندخل المتغير X_{3t} في نموذجنا لأن Y_t لا تعتمد عليه أي أن:

$$b_3 = 0$$

لاحظ (مع ذلك) أن هذا لا يمنعنا من كتابة نموذجنا على النحو:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + b_3 X_{3t} + u_t, \quad (6.107)$$

حيث إن $b_3 = 0$. تذكر أننا في مناقشتنا للملحق ب (B) للفصل الخامس، قد أوضحنا أنه طالما أن الصفر هو رقم ثابت، فلا يوجد، أساسا، خطأ في وجود معامل انحدار ذي قيمة مساوية للصفر.

افترض أننا لانعلم أن b_3 هي، في الحقيقة، مساوية للصفر. ونتيجة لذلك، نعد المعادلة (6.107) نموذجنا، ونكمل التحليل للحصول على مقدرات b_0, b_1, b_2 و b_3 بالطريقة العادية، حيثئذ، يكون السؤال الذي ينبغي أن نحاول الاجابة عنه هو: هل ستعطينا الطريقة العادية في التقدير مقدرات غير متحيزة لمعاملات الانحدار ولتباين مقدراتنا، والإجابة هي، لحسن الحظ، نعم، طالما أنه لم ينتهك أي من الافتراضات الأساسية المرتبطة بالخطأ العشوائي في المعادلة (6.107). افترض، على

سبيل المثال، أن u_t مستقلة عن X_{3t} . حيثئذ سيكون لدينا نموذج انحدار في المعادلة (6.107) تكون القيمة المتوسطة للخطأ العشوائي به مساوية للصفر، ويكون ذلك الخطأ العشوائي مستقلا عن المتغيرات المستقلة كافة كما يحقق جميع افتراضاتنا الأخرى. ويتبع عن ذلك أن مقدرات b_0, b_1, b_2, b_3 غير متحيزة.* وبديهيًا إذا لم نعرف أن تأثير X_{3t} على Y_t هو الصفر فإن البيانات ستدلتنا على ذلك، في شكل أن مقدرنا لـ b_3 سيكون غير متحيز $E(\hat{b}_3) = b_3 = 0$. وبالمقابل إذا أردنا اختبار فرضية العدم $b_3 = 0$ عند مستوى معنوية 5% فإنه ستكون لدينا فرصة 95% لقبوله.

تعقيبات إضافية

قد يبدو من النتائج السابقة أننا لانخسر شيئًا إذا ما أدخلنا مجموعة كبيرة من المتغيرات المستقلة المشكوك في أهميتها في المعادلة التي نرغب في تقديرها. ولكن هذا ليس صحيحًا. لاحظ أولاً أنه إذا كانت هذه التقديرات مستقلة عن الخطأ العشوائي حيث إنه لم ينتهك أي من افتراضاتنا فإنه لا يزال بإمكاننا رفض فرضية العدم المتعلقة بما إذا كانت إحدى الملمات المعنية تساوي الصفر أم لا. وبمعنى آخر، فقد تقع في الخطأ من النوع الأول.

ثانياً، وربما كان ذلك أكثر أهمية، يتزايد تباين مقدرات المعلمة - مع ثبات حجم العينة - عموماً - مع تزايد عدد المتغيرات المستقلة. أي أنه إذا علمنا أن $b_3 = 0$ وهكذا حصلنا على مقدر مثلاً لـ b_2 من المعادلة (6.106)، فإن هذا المقدر، عموماً، سيكون له تباين أصغر من مقدر b_2 الذي نحصل عليه من المعادلة (6.107) ويعني هذا أننا باستخدام المعادلة (6.107) نطلب كثيراً من البيانات المتاحة، طالما أنه

* هذه هي الحالة العادية التي يهتم بها في محاضرات الاقتصاد القياسي، ويمكن، في الحقيقة، إثبات أنه، إذا كانت X_{3t} غير مرتبطة بالخطأ العشوائي فإن مقدراتنا - مع توافر افتراضات إضافية - تظل متسقة. وأخيراً، ينبغي أن يكون واضحاً، أنه إذا كانت X_{3t} مرتبطة بالخطأ العشوائي، فإن مقدراتنا ستكون متحيزة وغير متسقة.

ينبغي أن نقدر أربع معلمات بدلاً من ثلاث. وبالمقابل، يدمج نموذج (6.106)، على العكس من (6.107) المعلومة بأن $b_3 = 0$ ، ويعكس تباين المقدرات هذه المعلومة الإضافية. إذا ثمة تكلفة مرتبطة مع زيادة عدد المتغيرات المستقلة: تباينات أكبر تقود لمقدرات أقل دقة للمعاملات. ونتيجة لذلك، تصبح فترات الثقة أوسع، وقد نرفض أحد المتغيرات المستقلة في النموذج على أساس أنه غير مهم إحصائياً بينما له في الحقيقة تأثير منتظم على المتغير التابع.

ويوقعنا هذا في معضلة. إذا تركنا أحد المتغيرات نحصل على نتائج متحيزة، ولكن، من ناحية أخرى، إذا أدخلنا عدداً كبيراً من المتغيرات في النموذج يزيد تباين مقدراتنا. ويشير هذا إلى الحاجة للاهتمام والاجتهاد عند اختيار مجموعة المتغيرات المستقلة. وحتى إذا كنا مهتمين، فقط، بالعلاقة بين متغيرين اثنين فقط (كما في المثال الموجود في الفصل الرابع حيث يهتم الباحث بتقدير تأثير معدل ضريبة المبيعات على مبيعات التجزئة في مراكز المدن)، فإنه من الضروري جعل النموذج يشتمل على متغيرات أخرى تحدد قيمة المتغير التابع. والفشل في ادخال المتغيرات المستقلة الأخرى سوف ينتج عنه، عموماً، مقدر متحيز للمعامل الذي نهتم به. وبالعودة مرة أخرى إلى ذلك المثال، سوف يتجه التأثير المقدر لمعدل الضرائب الأعلى على حجم المبيعات بمركز المدينة للتحيز إذا لم ندخل المتغيرات الأخرى المحددة لمبيعات التجزئة في المدينة. ومن الناحية الأخرى، لا يوجد سبب لادخال أي متغير تتوافر عنه البيانات باعتباره متغيراً مستقلاً في النموذج، ينبغي عليك أن تختار، فقط، تلك المتغيرات التي تعتقد على أسس مسبقة أنها تؤثر في المتغير التابع.

وأحد المناهج التي يستخدمها الاقتصاديون بانتظام هي أن «يجربوا» المتغير الذي يعتقدون بأهميته ثم يسقطونه من النموذج إذا تبين أنه ليس مختلفاً عن الصفر معنوياً. افترض، على سبيل المثال، أننا نريد تقدير المعادلة:

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t, \quad (6.108)$$

حيث تتوافر لدينا، فقط، أسباب واهية للاعتقاد بأن X_2 يؤثر في Y . ووجدنا فعلا أنه لا يمكننا رفض فرضية العدم بأن $b_2 = 0$. وأحد المناهج البديهية المغرية التي يمكن اللجوء إليها لتقليل تباين مقدرنا لـ b_1 هو إسقاط X_2 من المعادلة وتقدير المعادلة المنقحة:

$$Y_t = b_0 + b_1 X_{1t} + u_t, \quad (6.109)$$

وبينما يكون لهذا الأسلوب الشائع الاستخدام قية عملية بإسقاطه للمتغيرات التي ليست لها أهمية واضحة، فإننا يجب أن نكون على حذر من أن هذه الطريقة التي نستخدم، عادة، تتضمن تعقيدات معينة. وعلى سبيل أحد الأمثلة إذا كانت المجموعة الأولية من البيانات «معاداة استخدامها» لتقدير (6.109)، فإن عنصرا من الدائرية يدخل في النموذج، ويمكن اظهار أن المقدرات الناتجة تكون متحيزة. ويدخل هذا العنصر من الدائرية لأن النموذج (6.109) يعاد بناؤه على أساس الاختبار الذي أجرى على البيانات المستخدمة في تقدير النموذج. ومن الواضح أن الطريقة العلمية تتطلب من الباحثين أن يكونوا نماذجهم قبل تحليل البيانات التي ستستخدم لتقديرها.

هناك مشاكل أخرى ترتبط بهذه الطريقة المتابعة، وتظهر بعض هذه المشاكل حتى مع وجود مجموعة جديدة من البيانات تستخدم في تقدير (6.109). والمناقشة الكاملة لهذه المشاكل تقع خارج نطاق هذا الكتاب، ولكن ينبغي على القارئ أن يدرك أنه تنشأ مشاكل معينة كلما استخدمت نتيجة اختبار، في اختيار النموذج الذي يستخدم لتقدير المعلمة. وهذا ما فعلناه بالضبط عندما قدرنا b_1 بدلالة (6.108) إذا كان X_2 احصائيا معنوياً، وبدلالة (6.109) إذا كان X_2 غير معنوي. في مثل هذه الحالات، لا تحتفظ المقدرات بخصائصها المعتادة، فعلى سبيل المثال، قد تكون المقدرات متحيزة، وتبايناتها قد لا يمكن الحصول عليها بواسطة الصيغ النمطية، وقد لا تصبح موزعة توزيعاً طبيعياً. وعلى الرغم من أن هذه الطريقة المتابعة يتم استخدامها بصورة كبيرة في الاقتصاد، فإن المشاكل المترتبة عليها عادة ما يتم تجاهلها لسوء الحظ.

ملحق: ملاحظة حول الاستقرار

افترضنا في المعادلة (6.60) أن $|a| < 1$ ، وسبب ذلك هو أن الاختبارات المقترحة للارتباط الذاتي في الكتاب لن تكون صحيحة إلا إذا كانت $|a| < 1$ ، سنين في هذا الملحق أهمية هذا الافتراض.

بالإشارة إلى (6.60)، دع:

$$Z_t = b_0 + b_1 X_{1t} + \dots + b_k X_{kt} + u_t, \quad t = \dots, -2, -1, 0, 1, 2, \dots \quad (6A.1)$$

من الواضح أن Z_t هي صافي مجموع مكونات المتغيرات المستقلة فيما عدا aY_{t-1} . والآن يمكننا أن نعبر عن Y_t على النحو:

$$Y_t = aY_{t-1} + Z_t, \quad t = \dots, -2, -1, 0, 1, 2, \dots \quad (6A.2)$$

تتضمن العلاقة (6A.2)، أن:

$$Y_{t-1} = aY_{t-2} + Z_{t-1}. \quad (6A.3)$$

وبالتعويض عن (6A.3)، في (6A.2)، نحصل على:

$$Y_t = a^2 Y_{t-2} + aZ_{t-1} + Z_t. \quad (6A.4)$$

كما تتضمن العلاقة في (6A.2)، أيضا أن:

$$Y_{t-2} = aY_{t-3} + Z_{t-2}. \quad (6A.5)$$

وبالتعويض عن (6A.5)، في (6A.4)، نحصل على:

$$Y_t = a^3 Y_{t-3} + a^2 Z_{t-2} + aZ_{t-1} + Z_t. \quad (6A.6)$$

لاحظ أنه، في كل من (6A.2) و (6A.4) وأخيرا (6A.6) يعبر عن Y_t بالقيمة المبطأة لها والمضروبة في المعلمة a والمرفوعة، بدورها، إلى قوة معينة. وتعاود هذه القوة المرفوعة إليها a فترة الابطاء في Y_t . وهكذا، ففي (6A.6) مثلا نجد أن Y_t مبطأ ثلاث فترات، و a مرفوعة للقوة 3. وترتبط Y_t بكل من Z_t, Z_{t-1} و Z_{t-2} بطريقة مشابهة.

والآن، بأخذ قيمة Y_t في الحسبان والمبطأة t من الفترات، أي $Y_{t-t} = Y_0$.

وبالاستمرار في التعويض والذي يؤدي إلى (6A.6)، يمكننا التعبير عن Y_t بدلالة Y_0

على النحو:

$$Y_t = a^t Y_0 + a^{t-1} Z_1 + a^{t-2} Z_2 + \dots + a Z_{t-1} + Z_t. \quad (6A.7)$$

وعلى سبيل التوضيح، فإن (6A.7) تعني أن:

$$\begin{aligned} Y_1 &= aY_0 + Z_1 \\ Y_2 &= a^2 Y_0 + aZ_1 + Z_2 \\ Y_3 &= a^3 Y_0 + a^2 Z_1 + aZ_2 + Z_3, \text{ etc.} \end{aligned} \quad (6A.8)$$

دعنا الآن نهتم بالحالة التي يكون فيها $|a| < 1$. نجد هنا، أن اعتماد Y_t على قيمته المبطة t فترات سابقة (أي Y_0) تتضمن a^t التي تتناقص قيمتها المطلقة مع تزايد t ويعني هذا (من بين أشياء أخرى) أن اعتماد Y_2 على Y_0 سوف يكون أكبر في قيمته المطلقة عن اعتماد Y_3 على Y_0 . وبالمثل، فإن اعتماد Y_3 على Y_0 سوف يكون أكبر في قيمته المطلقة عن اعتماد Y_4 على Y_0 وهلم جرا.

فإذا ما تحققت شروط فنية معينة مرتبطة بمكونات Z_t^* ، واستمرت عملية الإحلال والتعويض (التي أوصلتنا إلى (6A.7) لانهائيا فسوف نحصل (في حالة $|a| < 1$) على:

$$Y_t = Z_t + aZ_{t-1} + a^2 Z_{t-2} + a^3 Z_{t-3} + \dots \quad (6A.9)$$

ويلاحظ أننا لم ندخل القيم المبطة لـ Y_t في الجانب الأيمن من (6A.9) لأن معامله سيكون صفرا إذا كانت $|a| < 1$.

افترض الآن الحالة التي تكون فيها $|a| \geq 1$ في هذه الحال، تصبح المناقشة السابقة التي تؤدي إلى (6A.9) غير صحيحة، وذلك طالما أن a^t لن يتناقص في قيمته المطلقة مع زيادة t . وبالمثل، نجد أنه، في ضوء (6A.7) و (6A.8)، يكون اعتماد Y_3 على Y_0 كبيرا في الأقل، كأعتماد Y على Y_0 . وبالمثل نفسه فإن اعتماد Y_4 على Y_0 سيكون كبيرا على الأقل كأعتماد Y_3 على Y_0 . وفي الحقيقة إذا كانت $|a| < 1$

* تتضمن هذه الشروط، حدسيا، أن احتمال أن تكون القيم المبطة لـ Z_t لانهائية هو الصفر.

** في عملية التعويض هذه نستخدم العلاقات $Y_{t-1} = aY_{t-2} + Z_{t-1}$ ، $Y_0 = aY_{-1} + Z_0$ وهلم جرا.

فإن اعتماد Y_t على Y_0 بالقيمة المطلقة سوف يتزايد كلما تزايدت t . وينبغي أن يكون ذلك واضحاً من (6A.7) و (6A.8).

دع Y_t هي، عموماً، قيمة المتغير التابع لنموذج معين في الفترة t . دع Y_0 هي قيمة ذلك المتغير في الفترة صفر. حيثُ إذا تضمن النموذج أن اعتماد Y_t على Y_0 يتناقص حتى يصل إلى الصفر عندما تؤول t إلى مالانهاية، فإن ذلك النموذج يكون مستقراً. أما إذا كان ذلك الاعتماد لا يتناقص إلى الصفر عندما تؤول t إلى مالانهاية، يكون النموذج غير مستقر. فإذا كان النموذج غير مستقر فإن قيمة المتغير التابع في أحد الفترات سوف تؤثر بصورة دائمة على قيمته المستقبلية.

تبين مناقشتنا أعلاه أن النموذج الموجود في المعادلة (6.60) هو نموذج مستقر إذا كانت $|a| < 1^*$ ، وغير مستقر إذا كانت $|a| > 1$. لاحظ أيضاً أن شروط الاستقرار في المعادلة (6.60) لا ترتبط بالمعاملات b_0, b_1, \dots, b_k .

أُسْئَلَة

١- افترض النموذج:

$$Y_t = a + bX_t + u_t,$$

مع المشاهدات

x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Y	2	2	2	1	3	5	6	6	10	10	10	12	15	10	11

باستخدام خطأ من النوع الأول قدره $\alpha = 0.05$ ، اختبر وجود الارتباط الذاتي، افترض تحقق الشروط المعتادة للانحدار.

* نفترض هنا أنها تحقق الشروط الفنية السابق ذكرها والمرتبطة بالقيمة المتباطئة لـ Z_t .

٢- افترض دالة الإنتاج الخطية التالية:

$$Q_t = a + bL_t + cK_t + u_t,$$

حيث يكون الناتج الكلي للاقتصاد في الفترة t ، Q_t ، مرتبطاً بكل من المدخلات الكلية للعمل ورأس المال L_t و K_t .

(أ) هل يمكن إثبات أن u_t في النموذج أعلاه يتسم باختلاف التباين؟ ناقش.

(ب) افترض أنه تم تجاهل رأس المال، وأصبح النموذج الذي قدر هو $Q_t = a + bL_t + u_t$. وضح لماذا يكون مقدر b متحيزاً (على الأرجح) لأعلى.

٣- افترض أن دالة الاستهلاك لكل فرد i ، في الفترة t تأخذ الشكل:

$$C_{it} = a + b_1 Y_{it} + b_2 Y_{it}^2 + u_{it}, \quad i=1, \dots, N. \quad (1)$$

اجعل C_t و Y_t متوسطي الإنفاق الاستهلاكي والدخل في الفترة t على الترتيب، أي:

$$C_t = \sum_{i=1}^N \frac{C_{it}}{N}, \quad Y_t = \sum_{i=1}^N \frac{Y_{it}}{N}. \quad (2)$$

حيث، في ضوء معادلات الاستهلاك الفردية التي تأخذ الشكل (١)، فإنه قد يمكننا أن نكون نموذج الانحدار الكلي التالي:

$$C_t = a + b_1 Y_t + b_2 Y_t^2 + u_t, \quad t=1, \dots, T. \quad (3)$$

(أ) بين أن النموذج أعلاه سيعاني خطأ التحديد (specification error).

يمكننا أن نطلق على هذا الخطأ «تحيز التجميع» طالما أنه ينشأ بسبب «تجميع» العلاقات الجزئية.

(ب) افترض أن $b_2 = 0$ ولذلك لا يوجد تحيز التجميع. افترض أنه يتوافر

لدينا مشاهدات عن $N=3$ أفراد لعدد $T=2$ سنوات. اكتب، بالرموز، مصفوفة المشاهدات التي تناظر المعادلة (١).

(ج) وضح لماذا سيكون لدينا، عمومًا، تحيز التجميع، إذا اشتمل نموذجنا للانحدار المقطعي cross-sectional على متغير غير خطي مثل Y_{it}^2 أعلاه؟
 -٤ افترض معادلة الطلب على النقود التالية:

$$M_{dt} = b_0 + b_1 i_t + b_2 i_{t-1} + b_3 (\Delta i_t) + u_t,$$

حيث إن M_{dt} : الطلب على النقود، i_t معدل الفائدة و $i_t - i_{t-1} = \Delta i_t$. نفترض أن i_{t-1} تعكس تأثير العادة، أو القصور الذاتي، وأن Δi_t يعكس «أثر التوقع» الناتج عن التغيرات الحديثة في معدلات الفائدة.
 (أ) أثبت، بدون معلومات إضافية، أنه من غير الممكن تقدير أي من b_1, b_2, b_3 .

(ب) ضع المعادلة في شكل يمكن تقديره، ويمكن استخدامه للتنبؤ بقيم M_{dt} .

-٥ افترض دالة الإنتاج التالية:

$$Q_t = A L_t^{\alpha_1} L_2^{\alpha_2} K_t^{\alpha_3} e^{u_t},$$

حيث L_1 = عدد عمال الإنتاج، L_2 = عدد العمال الآخرين، k = رصيد رأس المال، u_t = الخطأ العشوائي، ويشير الرمز t إلى الفترات الزمنية. افترض أن المنشأة توظف دائمًا 10,000 عامل. حيث $L_{1t} + L_{2t} = 10,000$ ، هل يمكننا تقدير α_1, α_2 و α_3 ؟ وضح.

-٦ افترض النموذج التالي:

$$Y_t = a + bX_t + u_t,$$

$$u_t = \rho u_{t-1} + \varepsilon_t.$$

افترض أننا نحسب \hat{b} بوساطة الصيغة المعتادة:

$$\hat{b} = \frac{\sum (X_t - \bar{X})(Y_t - \bar{Y})}{\sum (X_t - \bar{X})^2}$$

بين أنه، بسبب الارتباط الذاتي، لم تعد صيغتنا المعتادة لتباين \hat{b} التي تأخذ الشكل:

$$\sigma_{\hat{b}}^2 = \frac{\sigma_u^2}{\sum (X - \bar{X})^2}$$

صحيحة.

٧- إجعل دالة الاستهلاك تأخذ الشكل:

$$C_t = b_0 + b_1 Y_t + b_2 A_t + u_t,$$

حيث يفترض بالنسبة لكل قيمة من قيم Y_t و A_t أن $E(u_t) = 0$ وتباين u_t هو $Y_t^2 \sigma_u^2$. حول المعادلة السابقة إلى معادلة أخرى يتسم فيها الخطأ العشوائي بثبات التباين، ثم اشتق بعد ذلك المعادلات الطبيعية.

٨- افترض النموذج التالي:

$$I_t = a_0 + a_1 \Delta Y_t + a_2 r_t + \varepsilon_t,$$

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + u_t,$$

حيث I = الاستثمار و ΔY = التغير في الدخل و r_t معدل الفائدة. افترض أن u_t مستقل عن المتغيرات المستقلة كافة، ولا يتسم بوجود الارتباط الذاتي، وله قيمة متوسطة صفرية، وتباين ثابت. وضح باختصار الطريقة الذي يمكننا اتباعه للحصول على تقديرات a_0 ، a_1 و a_2 والتي تأخذه في الحسبان مشكلة الارتباط الذاتي عندما يكون كل من ρ_1 و ρ_2 غير معلومين.

نظم المعادلات

درسنا حتى الآن التقدير لمعادلة واحدة بمعزل عن النموذج الاقتصادي الأكبر الذي قد تكون هذه المعادلة جزءاً منه. على سبيل المثال تكون معادلة الطلب على سلعة معينة، عادة، معادلة واحدة ضمن نظام من المعادلات يحدد السعر التوازني وكمية تلك السلعة في السوق، وعموماً، يشتمل النموذج الاقتصادي لسوق معين على معادلة للطلب ومعادلة للعرض، وأخيراً، على معادلة ثالثة تصف العملية التوازنية في السوق (أي مساواة الكمية المطلوبة بالكمية المعروضة). سنهتم في هذا الفصل بالمشاكل التي قد تنشأ عندما تكون المعادلة المراد تقديرها مرتبطة بمعادلات أخرى في نموذجنا الأكبر. وسنجد بخاصة أنه، في ظل ظروف معينة، لم يعد بإمكان طريقتنا المعتادة في التقدير إعطاءنا مقدرات غير متحيزة (أو حتى متسقة) للمعاملات. وفي هذه الحالات، علينا أن نعدل من طريقتنا للتقدير.

(٧-١) تحيز المعادلات الآتية

نذكر أننا، عند مناقشتنا الأولية لنموذج الانحدار في الفصل الثاني، كوّنّا فرضين مرتبطين بخصائص معادلة الانحدار:

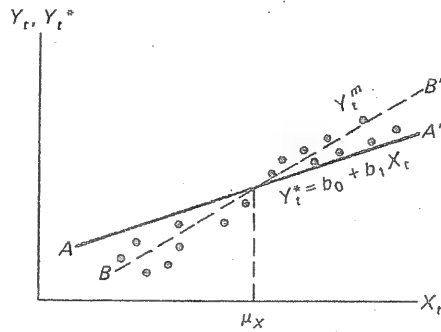
$$Y_t = b_0 + b_1 X_t + u_t.$$

فقد فرضنا أن القيمة المتوسطة للخطأ العشوائي هي الصفر $E(u_t) = 0$ وأن الخطأ العشوائي مستقل عن المتغير المستقل في النموذج، ولذا يكون التباين بينهما مساوياً

للصفر $E(X_t, u_t) = 0$. وحيث، أمكننا وفقا لهذه الافتراضات، أن نضع في طريقتنا للتقدير:

$$\sum \hat{u}_t = 0, \quad \sum (X_t \hat{u}_t) = 0.$$

قد مكنتنا ذلك من الحصول على معادلتين طبيعيتين، وبحلها معا، حصلنا على مقدرات معلماتنا \hat{b}_0, \hat{b}_1 ، وحيث، أمكننا إثبات أن هذه المقدرات غير متحيزة. لاحظنا، أيضا، أن غياب التحيز في هذه المقدرات يعتمد على صحة افتراضاتنا. وعلى سبيل المراجعة افترض أن التغير بين X_t ، u_t لم يكن صفرا أي $\text{cov}(X_t, u_t) = E(X_t u_t) \neq 0$. افترض، على سبيل المثال، أن التغير موجب. يعني هذا أن القيم الأكبر من القيمة المتوسطة لـ u_t (وهي قيم موجبة طالما أن متوسط u_t يساوي صفرا) سوف تكون مرتبطة بقيم أكبر من القيمة المتوسطة لـ X_t والعكس بالعكس. يتضمن هذا، بدوره، أن القيمة المتوسطة لـ Y_t سوف تكون أكبر من $Y^* = (b_0 + b_1 X_t)$ عندما تكون X_t أكبر من قيمتها المتوسطة، μ_x ، وأقل من Y^* عندما تكون X_t أقل من μ_x . افترض، في شكل الانتشار (٧-١) أن العلاقة المتضمنة لـ Y_t^* ممثلة بالخط AA' . حيث، طالما أن القيم الموجبة لـ u سوف تحدث عادة عندما تكون X كبيرة، في حين تحدث القيم السالبة لـ u مصاحبة للقيم الأصغر من X ، إن الانتشار المشاهد للنقط يقع حول منحني يشبه BB' . فإذا ما افترضنا على نحو غير صحيح أن $E(X_t u_t) = 0$ ، وتبعنا لذلك، وضعنا الفرض $\sum X_t \hat{u}_t = 0$ ، فسوف نصل في النهاية إلى تقدير معلمات العلاقة BB' . بمعنى أننا سوف ننتهي بتقدير العلاقة بين Y_t ، X_t التي يتوقع أن تكون في منتصف جميع النقاط في شكل الانتشار. وسينتج عن ذلك مقدرات متحيزة للمعلمات b_0 و b_1 ، وفي هذه الحال، سنميل إلى بخس تقدير b_0 (الجزء المقطوع من المحور الرأسي لـ AA' والمبالغة في تقدير b_1 (ميل AA'). ونؤكد هنا أن هذا التحيز لا تقل أهميته عند كبر حجم العينة. وهنا لن تكون مقدراتنا متحيزة، فقط، بل غير متسقة أيضا.



شكل رقم (٧-١)

دعنا الآن نعود إلى دالتنا البسيطة للاستهلاك والتي تأخذ الشكل *:

$$C_t = b_0 + b_1 Y_t + u_t. \quad (7.1)$$

في هذه المعادلة، نفترض أن الخطأ العشوائي u_t موزع توزيعاً طبيعياً وله قيمة متوسطة صفرية، $E(u_t) = 0$ ، وتباين ثابت $E(u_t^2) = \sigma_u^2$ ، وغير مرتبط ذاتياً. لا يفترض

هنا استقلال u_t عن Y_t أو أنها غير مرتبطة بها، للأسباب التي سنناقشها الآن: نعلم من النظرية الاقتصادية الكلية أنه توجد، في الأقل، معادلة إضافية

أخرى في هذا النموذج، وهي معادلة توازن تبين أن مستوى الناتج والدخل يستمر في التغير حتى يتعادل الطلب الكلي $(C_t + I_t)$ مع العرض الكلي Y_t :

$$Y_t = C_t + I_t, \quad (7.2)$$

حيث إن I_t هو مستوى الإنفاق الاستثماري بوساطة الأفراد ومؤسسات الأعمال**.

دعنا نفترض أن I_t متغير خارجي، يعني هذا أن قيمته في أي فترة زمنية t تحدد بوساطة عوامل خارج النموذج، فقد يتحدد الإنفاق الاستثماري، مثلاً، بوساطة

* لن نميز في النموذج المبسط الذي كوّناه في هذا البحث بين الدخل الكلي و الدخل المتاح، واتباعاً لما جرى عليه العرف في الكتب الأخرى، سنرمز للدخل بالرمز Y ونعد، ببساطة، أن الاستهلاك دالة في الدخل.

** تعد المعادلة (7.2) معادلة توازن بسيطة جداً لأنها تهمل طلب القطاع الحكومي وصافي الطلب الخارجي، ويمكننا، ببساطة، أن ندخل هذه المكونات الأخيرة للطلب في النموذج، ولكن ذلك غير ضروري للمشكلة التي نهتم بها هنا.

قوة العادة (حيث يزيد الاستثمار في إحدى الفترات بمقدار ٢٪ مثلاً عن الفترة السابقة لها) أو بعوامل اجتماعية. سنفترض في هذا النموذج، على أية حال، أن هذه «القوى الخارجية» مهما كان نوعها، ليست مرتبطة بالخطأ العشوائي u_t في (7.1) ولذا، فإن التغير بين I_t و u_t يكون صفراً.

$$E(I_t u_t) = \text{cov}(I_t, u_t) = 0.$$

سنبين الآن - ويسبب العلاقة بين Y_t و C_t في المعادلة (7.2) - أن التغير بين الدخل Y_t والخطأ العشوائي u_t في (7.1) ليس صفراً. ويمكننا توضيح ذلك بسهولة إذا ما قمنا بالتعويض عن قيمة C_t من (7.1) في (7.2) لنحصل على:

$$Y_t = b_0 + b_1 Y_t + u_t + I_t. \quad (7.3)$$

وبحل هذه المعادلة للحصول على Y_t ، نصل إلى:

$$Y_t = \frac{b_0}{1-b_1} + \frac{I_t}{1-b_1} + \frac{u_t}{1-b_1} \quad (7.4)$$

فإذا ما ضربنا، بعد ذلك، (7.4) في u_t وأخذنا القيم المتوقعة لحصلنا على:

$$\begin{aligned} \text{cov}(Y_t, u_t) &= E(Y_t u_t) = E\left(\frac{b_0 u_t}{1-b_1} + \frac{I_t u_t}{1-b_1} + \frac{u_t^2}{1-b_1}\right) \\ &= \frac{b_0}{1-b_1} E(u_t) + \frac{1}{1-b_1} E(I_t u_t) + \frac{1}{1-b_1} E(u_t^2). \end{aligned} \quad (7.5)$$

ولما كانت القيمة المتوسطة لـ u_t هي الصفر وافترضنا أن التغير بين u_t و I_t مساو للصفر، فإن الحدين الأولين في الصيغة الأخيرة من (7.5) يساويان الصفر، ولكن الحد الأخير ليس صفراً. وبالتحديد، يكون لدينا:

* نذكر، مرة أخرى، أنه طالما أن القيمة المتوسطة لـ u_t هي الصفر، فإن التغير بين u_t و Y_t هو $E(u_t Y_t)$

$$\begin{aligned} E[(u_t - 0)(Y_t - \mu_Y)] &= E(u_t Y_t) - E(u_t \mu_Y) \\ &= E(u_t Y_t) - \mu_Y E(u_t) = E(u_t Y_t), \end{aligned} \quad \text{لأن:}$$

حيث إن μ_Y هو القيمة المتوسطة لـ Y_t

$$E(Y_t u_t) = \frac{1}{1-b_1} E(u_t^2) = \frac{\sigma_u^2}{1-b_1} \neq 0. \quad (7.6)$$

نلاحظ أنه، وبسبب المعادلة الإضافية (7.2)، لم يعد بالإمكان افتراض أن التغير بين المتغير المستقل Y_t والخطأ العشوائي u_t في (7.1) هو الصفر. وتشير المعادلة (7.6) إلى أن واحدا من الافتراضات الأساسية لنموذجنا للانحدار لم يعد صحيحا، لذلك، فإن استخدام طريقتنا المعتادة للتقدير سوف يؤدي - لأسباب ذكرناها من قبل - إلى إيجاد مقدرات متحيزة وغير متسقة لكل من b_1 و b_0 .

قد يكون من المفيد الآن إكمال معالجتنا النظرية هذه بمناقشة أكثر بديهية لمصدر هذا التحيز. افترضنا في المعادلة (7.1) أن الإنفاق الاستهلاكي في الفترة C_t, t يعتمد على الدخل في الفترة الزمنية نفسها، Y_t . ولكننا نرى من (7.2) أن Y_t تعتمد، بدورها، على C_t . توجد لدينا إذا علاقة سببية تعمل في كلا الاتجاهين. فالمتغيران متداخلان: أي يعتمد كل منهما على الآخر. فإذا قمنا بتقدير (7.1) بواسطة طريقتنا المعتادة فإن ذلك يعني أننا نحدد نمط العلاقة الإحصائية بين C_t و Y_t . فإذا كانت العلاقة السببية تعمل في اتجاه Y_t إلى C_t بدون غموض أو لبس، يمكننا، حينئذ، أن تقديراتنا نفسها على أنها تشير إلى تأثير Y_t على C_t . ولكن، في حالة الاعتماد المتبادل [كما يظهر من (7.2)]، يعكس هذا الارتباط الإحصائي المحسوب بين C_t و Y_t خليطا من تأثير كل من المتغيرين على الآخر. وفي هذه الحال، لا يمكننا تفسير تقديراتنا على أنها مقياس لا يتسم بالغموض لتأثير أحد المتغيرات، وهنا Y_t ، على الآخر، C_t . ونحتاج هنا إلى طريقة منقحة للتقدير تسمح لنا بفصل الأثرين عن بعضهما بعضا، أي إلى طريقة لعزل تأثير Y_t على C_t من تأثير C_t على Y_t .

هذه هي مشكلة تحيز المعادلات الآتية. وتنشأ هذه المشكلة، عموما، عندما تكون قيمة أحد المتغيرات المستقلة ذاتها دالة في المتغير التابع. افترض نموذج الانحدار المتعدد التالي :

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_t, \quad (7.7)$$

حيث X_{1t} هو المتغير المستقل الذي يحقق شروطنا المعتادة من حيث إنه مستقل عن الخطأ العشوائي. فإذا كان صحيحا أن X_{2t} يعتمد على Y_t ، حينئذ، تكون لدينا معادلة أخرى للاهتمام بها - ربما تأخذ الشكل :

$$X_{2t} = c_0 + c_1 Y_t + r_t, \quad (7.8)$$

حيث إن r_t هو الخطأ العشوائي الذي يفترض، أيضا، استقلاله عن X_{1t} . سنترك للقارئ، على سبيل التدريب، أن يثبت (كما فعلنا من قبل) أنه

$$E(X_{2t} u_t) \neq 0.$$

تتضمن هذه المجموعة من المعادلات انتهاكا لواحد من الافتراضات الأساسية لنموذج الانحدار، وأي محاولة لاستخدام منهجنا المعتاد لتقدير (7.7) سوف ينتج عنه مقدرات متحيزة وغير متسقة للمعاملات كافة في المعادلة.

سوف نشير، من الآن فصاعدا، لخاصية الاتساق أو عدمها، فقط، لمقدراتنا. وسبب ذلك هو أن الاقتصاديين القياسيين يحلون «مشكلة نظم المعادلات الآتية» عن طريق بناء مقدرات متسقة لأنه لا توجد، عموما، مقدرات غير متحيزة للمعاملات في المعادلات التي تكون جزءا من نظام أكبر من المعادلات.

(٧-٢) طريقة المربعات الصغرى ذات المرحلتين : حالة مبسطة

تمكن الاقتصاديون القياسيون في السنوات الأخيرة من تطوير عدد من الطرق المختلفة لمعالجة مشاكل التقدير الخاصة بنظم المعادلات الآتية*، سوف نبين أحدها هنا: طريقة المربعات الصغرى ذات المرحلتين (م ص م) Two-stage least squares. وتتسم هذه الطريقة بعدد من السمات الجيدة. فهي أولا، طريقة يسهل

* كما سنرى بعد قليل، فإن بعض المعادلات، أو معاملات معينة بها، لا يمكن تقديرها. وفي الحقيقة، فإننا قد واجهنا هذه المشكلة في مجال آخر، وهو مجال الارتباط الخطي المتعدد، حيث لم نستطع تقدير بعض المعاملات في بعض المعادلات.

فهمها وتعتمد اعتمادا كبيرا على المادة العلمية التي تناولناها في هذا الكتاب، ثانيا، وفي ظل تحقق الشروط المعتادة، إذا كان يمكن تقدير المعادلة موضع الاهتمام، فإن م ص م (وعلى العكس من طرق التقدير الأخرى) تعمل دائما حيث توجد مقدرات متسقة للمعلومات الممكن تقديرها. ثالثا، يمكن أن نطلق على م ص م بأنها طريقة المعلومات المحدودة Limited-information. ونعني بذلك أنه يمكن تقدير معادلة معينة موجودة في إطار معادلات آنية عن طريق م ص م باستخدام معلومات عامة غير محددة، فقط، عن المعادلات الأخرى. بينما يتطلب تنفيذ عديد من الطرق الأخرى معلومات أكثر تفصيلا عن تلك المعادلات. رابعا، تحتاج م ص م إلى قدر متواضع من الحسابات مقارنة بالطرق الأخرى.

توضيح : المقدرات المتسقة

سنقدم أولا مقدمة حدسية لم ص م. نذكر أن مشكلتنا، أساسا، هي وجود العلاقة السببية ذات الاتجاهين التي نجم عنها تغاير لا يساوي الصفر بين الخطأ العشوائي وواحد أو أكثر من المتغيرات المستقلة. فإذا أردنا استخدام طريقة المربعات الصغرى العادية في التقدير (م ص م)*، فإنه ينبغي علينا - بطريقة أو بأخرى - التخلص من التغاير غير الصفري حتى تستوفي المعادلة المقدرة افتراضات نموذجنا للانحدار. وهذا ما تفعله، بالضبط، م ص م. إنها طريقة ذات مرحلتين، في المرحلة الأولى، نزيل من المتغير (أو المتغيرات) المستقلة ذلك الجزء المرتبط بالخطأ العشوائي. ويتضمن ذلك إيجاد مجموعة منقحة من القيم للمتغيرات المستقلة المشكوك فيها. هذه القيم المنقحة لا تكون مرتبطة بالخطأ العشوائي، ولذا

* نذكر من الفصل الثاني أن طريقتنا للتقدير باستخدام المتغير المساعد لها سمة «المربعات الصغرى»، بمعنى أن تلك الطريقة تكون مماثلة لتدنية مجموع الانحرافات المربعة للنقاط المشاهدة عن خط الانحدار المقدّر. لهذا السبب، قد نشير إلى نتائجنا المعتادة على أنها نتائج طريقة المربعات الصغرى العادية (م ص ع)، ويعد هذا الترميز ملائماً للتمييز بين طريقتنا المعتادة م ص ع وطريقة م ص م.

تكون الخطوة الثانية، ببساطة، هي تقدير المعلمات بطريقتنا العادية للتقدير. ولنعرف كيف يتم ذلك، دعنا نعود إلى نموذجنا المبسط ذي المعادلتين للدخل القومي حيث لدينا :

$$C_t = b_0 + b_1 Y_t + u_t, \quad (7.1)$$

$$Y_t = C_t + I_t. \quad (7.2)$$

وكما سبق، فإنه، بالتعويض عن C_t من (7.1) في (7.2) وحلها للحصول على Y_t نحصل على :

$$Y_t = \frac{b_0}{1-b_1} + \frac{I_t}{1-b_1} + \frac{u_t}{1-b_1}. \quad (7.4)$$

ويمكننا أن نبسط رموزنا عن طريق كتابة (7.4) على النحو :

$$Y_t = a_0 + a_1 I_t + q_t, \quad (7.9)$$

حيث :

$$q_t = \frac{u_t}{1-b_1}, \quad a_1 = \frac{1}{1-b_1}, \quad a_0 = \frac{b_0}{1-b_1},$$

لدينا الآن المتغير Y_t بوصفه دالة في I_t وفي الخطأ العشوائي q_t . لاحظ أن q_t (مثل u_t) له متوسط يساوي الصفر. افترض للحظة أننا نعلم قيم a_0 و a_1 . في ظل هذا الافتراض يمكننا اشتقاق متوسط Y_t ، أي Y_t^m المرتبط بأي قيمة معطاة لـ I_t :

$$Y_t^m = a_0 + a_1 I_t. \quad (7.10)$$

ومن (7.9)، يمكننا أيضا، رؤية أن :

$$Y_t = Y_t^m + q_t. \quad (7.11)$$

وبالرجوع إلى دالتنا الأصلية للاستهلاك (7.1)، وبالتعويض في (7.11) عن قيمة Y_t ، نحصل على :

$$\begin{aligned} C_t &= b_0 + b_1(Y_t^m + q_t) + u_t \\ &= b_0 + b_1 Y_t^m + (b_1 q_t + u_t). \end{aligned} \quad (7.12)$$

وبملاحظة أن $q_t = u_t / (1 - b_1)$ ، يمكننا كتابة (7.12) على النحو التالي :

$$\begin{aligned} C_t &= b_0 + b_1 Y_t^m + \frac{u_t}{1 - b_1} \\ &= b_0 + b_1 Y_t^m + q_t. \end{aligned} \quad (7.13)$$

سنبين الآن أن المعادلة (7.13)، بعكس (7.1)، تحقق الافتراضات الأساسية لنموذجنا للانحدار. فأولا، يمكننا رؤية أن الخطأ العشوائي له متوسط يساوي الصفر.

$$E(q_t) = E\left(\frac{u_t}{1 - b_1}\right) = \frac{1}{1 - b_1} E(u_t) = 0.$$

وثانيا، نجد أن المتغير المستقل لم يعد مرتبطا بالخطأ العشوائي، فإذا ضربنا (7.10) بـ q_t وأخذنا القيمة المتوقعة، نحصل على :

$$\begin{aligned} E(Y_t^m q_t) &= E(a_0 q_t + a_1 I_t q_t) \\ &= \left(\frac{a_0}{1 - b_1}\right) E(u_t) + \left(\frac{a_1}{1 - b_1}\right) E(I_t u_t) = 0, \end{aligned} \quad (7.14)$$

وطالما أن القيمة المتوقعة لـ u_t تساوي صفرا، وطالما أن I_t مستقل عن u_t افتراضا*، لذا، يمكننا استخدام طريقتنا المعتادة في التقدير للحصول على مقدرات متسقة لـ b_0 و b_1 . وسنفرض الشرطين $\Sigma \hat{q}_t = 0$ و $\Sigma Y_t^m \hat{q}_t = 0$ لاشتقاق المعادلتين الطبيعيين :

* من الواضح أن الخطأ العشوائي q_t لن يعاني مشكلة الارتباط الذاتي (طالما أن ليس مرتبطا ذاتيا)، وسيكون له تباين ثابت، هو $[\sigma_u^2 / (1 - b)^2]$. يضاف إلى ذلك أن q_t سيكون موزعا توزيعا طبيعيا طالما أن u_t موزع كذلك.

$$\sum C_t = n\hat{b}_0 + \hat{b}_1 \sum Y_t^m,$$

$$\sum (C_t Y_t^m) = \hat{b}_0 \sum Y_t^m + \sum (Y_t^m)^2,$$

واللتين يمكننا حلهما للحصول على \hat{b}_0 و \hat{b}_1 . لاحظ، أيضا، أن هاتين المعادلتين هما اللتان سنحصل عليهما لو أحللنا Y_t^m محل Y_t في المعادلة الأصلية (7.1) ومن ثم، واصلنا لنشتق المعادلتين الطبيعيتين بالطريقة المعتادة.

وعلى سبيل المراجعة، وللتعامل مع مشكلة نظم المعادلات، نحدد أولا القيمة المتوسطة للدخل Y_t^m المرتبطة بكل قيمة للاستثمار، I_t . أما الخطوة الثانية فتمثل بإحلال هذا المتغير الجديد محل Y_t في المعادلة الأصلية، ومن ثم، تقدير المعلومات بالطريقة المعتادة. وبالطبع، عندما تحل Y_t^m محل Y_t في معادلتنا الأساسية (7.1)، فإننا نلاحظ أن الخطأ العشوائي، q_t ، في معادلتنا الناتجة (7.13) ليس ماثلا للخطأ العشوائي الأصلي، u_t ، وفي ضوء أهدافنا فهذا التغير غير مهم، في الحقيقة، وهو فعلا يرتبط بالترميز، فقط، لأننا بينا أن q_t لها السمات الأساسية نفسها كما لـ u_t .

وتتكون طريقة م ص م، في الواقع، من تنقية المتغير المستقل، Y_t ، من ذلك الجزء، q_t ، المرتبط بالخطأ العشوائي الأصلي u_t . تذكر أن $Y_t = Y_t^m - q_t$. وكما أوضحنا أن Y_t^m غير مرتبط بـ q_t (وبالتالي بـ u_t) وعندما نستخدم Y_t^m بدلا من Y_t فنحن بالفعل قد طرحنا q_t من Y_t : $Y_t^m = Y_t - q_t$ بمعنى أننا قد أزلنا ذلك الجزء من Y_t الذي يحتوي على تأثير الجزء العشوائي للمتغير التابع C_t على المتغير المستقل Y_t ، وبهذه الطريقة يمكننا، باستخدام م ص م، فصل التأثيرات التي كانت متداخلة من قبل.

وبينما تظهر هذه الطريقة من حيث المبدأ سليمة، فإن الصعوبة العملية تكمن في عدم امكانية تنفيذ ذلك مباشرة لأننا لانعرف قيم Y_t^m . تذكر أنه في المعادلة (7.10):

$$Y_t^m = a_0 + a_1 I_t. \quad (7.10)$$

افترضنا أن قيم كل من a_0 و a_1 معلومة. وهذا لن يكون صحيحا، عموما، إلا أنه يمكننا تقدير قيمتها. فإذا نظرنا إلى الخلف للمعادلة التي اشتققناها من النموذج الأصلي، أي (7.9)، يمكننا أن نتأكد أنها تحقق افتراضات نموذج الانحدار الأساسي كافة:

$$Y_t = a_0 + a_1 I_t + q_t. \quad (7.9)$$

وباستطاعتنا أن نحصل على مقدرات لـ a_0 و a_1 : \hat{a}_0 و \hat{a}_1 مثلا، من المعادلات الطبيعية التي نوجدتها عن طريق وضع $\Sigma \hat{q}_t = 0$ و $\Sigma(\hat{q}_t I_t) = 0$. وباستطاعتنا، حينئذ، استخدام \hat{a}_0 و \hat{a}_1 للحصول على مقدر لـ Y_t^m :

$$Y_t^m = \hat{a}_0 + \hat{a}_1 I_t. \quad (7.15)$$

وبالعودة إلى الرموز المستخدمة في الفصول السابقة، نجد أن Y_t^m ليست سوى القيمة المحسوبة لـ Y_t والمناظرة للانحدار (7.9) : $\hat{Y}_t = \hat{Y}_t^m$. والمرحلة الثانية هي استخدام \hat{Y}_t (بدلا من Y_t^m) لتقدير معادلة الاستهلاك، أي أننا سنجعل نموذجنا للانحدار يأخذ الشكل :

$$C_t = b_0 + b_1 \hat{Y}_t + u_t^*, \quad (7.16)$$

حيث إن $u_t^* = u_t + b_1 \hat{q}_t$ ، طالما أن $Y_t = \hat{Y}_t + \hat{q}_t$. وباتباع المنهج المعتاد سنفترض الشروط $\Sigma \hat{u}_t^* = 0$ و $\Sigma(\hat{u}_t^* \hat{Y}_t) = 0$ للحصول على المعادلات الطبيعية :

$$\begin{aligned} \Sigma C_t &= n\hat{b}_0 + \hat{b}_1 \Sigma \hat{Y}_t, \\ \Sigma(C_t \hat{Y}_t) &= \hat{b}_0 \Sigma \hat{Y}_t + \hat{b}_1 \Sigma \hat{Y}_t^2, \end{aligned}$$

التي سنقوم بحلها للحصول على \hat{b}_0 و \hat{b}_1^* . وفي الحقيقة يمكن إثبات (في ظل شروط عامة) أن كل من \hat{b}_0 و \hat{b}_1 التي حصلنا عليها متسقتان.

خلاصة القول لحالتنا التوضيحية هذه تتكون طريقة م ص م من أولاً: انحدار المتغير المستقل المشكوك فيه، Y_t ، على المتغير الخارجي، I_t ، واستخدام هذه المعادلة المقدرة لاشتقاق المتغير المستقل الجديد، \hat{Y}_t . وثانياً: إحلال \hat{Y}_t محل Y_t في معادلتنا الأصلية، وبعدئذ نقدر معادلتنا بالطريقة المعتادة. ونلاحظ هنا وفي ظل هذه الطريقة أن الخطأ العشوائي في المرحلة الثانية u_t^* هو تعديل مبسط للخطأ العشوائي الأصلي u_t .

بعض النتائج الإضافية

قبل المضي في تعميم طريقتنا، ينبغي أن نلاحظ أنه، عندما تكون لدينا مقدرات متسقة لـ b_0 و b_1 ، يمكننا وبساطة الحصول على مقدر متسق للخطأ العشوائي u_t في (7.1) بواسطة القاعدة الواضحة التالية:

$$\hat{u}_t = C_t - (\hat{b}_0 + \hat{b}_1 Y_t).$$

* من الناحية العلمية، لاحظ a_0 و a_1 في (7.9) سيتم تقديرها عن طريق فرض الشروط:

$$\Sigma \hat{q}_t = 0 \text{ و } \Sigma (\hat{q}_t I_t) = 0$$

وطالما أن $\hat{Y}_t = \hat{a}_0 + \hat{a}_1 I_t$ فإنه يتبع عن ذلك أن $\Sigma (\hat{q}_t \hat{Y}_t) = 0$ والآن، بسبب أن $u_t^* = u_t + b_1 \hat{q}_t$

في (7.16)، يكون لدينا:

$$\Sigma u_t^* = \Sigma \hat{u}_t = 0, \quad E(u_t) = 0,$$

وتفيد هذه الشروط أنه، لتقدير (7.16) نضع:

$$\Sigma \hat{u}_t^* = \Sigma \hat{u}_t = 0, \quad \text{فإن } E(u_t) = 0, \quad \text{وطالما}$$

$$\Sigma (\hat{u}_t^* \hat{Y}_t) = \Sigma (\hat{u}_t \hat{Y}_t) = 0, \quad \text{فإن } E(u_t Y_t) = 0. \quad \text{وطالما}$$

ويعني آخر لا يؤدي مكون \hat{q}_t في u_t^* أي دور في تقدير (7.16) وسنرى النتائج المترتبة على ذلك بالنسبة لصيغ التباين فيما بعد.

وبهذا، نحصل على مقدر متسق لتباين u_t بوساطة صيغتنا المعتادة :

$$\hat{\sigma}_u^2 = \frac{\sum \hat{u}_t^2}{n-2} \quad (7.17)$$

وأحد السمات الجيدة الأخرى لطريقة م ص م هو أنه متى تم إحلال Y_t^m محل Y_t فإن صيغنا القديمة لتباين \hat{b}_0 و \hat{b}_1 تظل صحيحة فيما عدا تغيرا في التفسير . وبالإشارة إلى (7.16) ، ويذكر أن متوسط العينة لـ \hat{Y}_t هو $\bar{Y}_t = \bar{Y}$ ، فإنه يمكن إثبات أن الصيغ المعتادة :

$$\sigma_{\hat{b}_0}^2 = \sigma_u^2 \left[\frac{\sum \hat{Y}_t^2}{n \sum (\hat{Y}_t - \bar{Y})^2} \right], \quad (7.18)$$

$$\sigma_{\hat{b}_1}^2 = \sigma_u^2 \left[\frac{1}{\sum (\hat{Y}_t - \bar{Y})^2} \right], \quad (7.19)$$

فإنه يمكن إثبات أنها مازال صحيحة بشرط أن تكون العينة كبيرة لا نهائية . وطالما أنه لا توجد لدينا عينات ذات حجم لانهائي ، فإنه ينبغي علينا ، مرة ثانية ، تفسير هذه الصيغ باعتبارها قيما تقريبية .

سوف يلاحظ القارئ الفطن اختلافا واضحا في صيغ التباين . وبالتحديد ، فطالما أن التقدير في المرحلة الثانية يتضمن المعادلة (7.16) حيث يصبح الخطأ العشوائي u_t^* (وليس u_t) ، فإن صيغنا لتباين \hat{b}_0 و \hat{b}_1 في (7.18) و (7.19) ينبغي أن تتضمن تباين u_t^* (أي $\sigma_{u^*}^2$) بدلا من σ_u^2 . ولكن الحال ليست كذلك وسبب ذلك هو أن جزء \hat{q}_t في صيغة u_t^* .

$$u_t^* = u_t + b_1 \hat{q}_t,$$

يلغي بعضه بعضا في كل من \hat{b}_0 و \hat{b}_1 . ويتبع عن ذلك أنه يمكن إثبات أن قيم كل من هذين المقدرين تعتمد على u_t (وليس على u_t^*) . ومن ثم ، يعتمد كل من \hat{b}_0 و \hat{b}_1 على تباين u_t وليس على تباين u_t^* .

ولنرى ذلك، لاحظ أن مقدارنا \hat{b}_1 سيكون :

$$\hat{b}_1 = \frac{\sum (\hat{Y}_t - \bar{Y}) C_t}{\sum (\hat{Y}_t - \bar{Y})^2} \quad (7.20)$$

وبالتعويض عن C_t من (7.16) في (7.20) وبالاختصار، نحصل على :

$$\hat{b}_1 = b_1 + \frac{\sum (\hat{Y}_t - \bar{Y}) u_t^*}{\sum (\hat{Y}_t - \bar{Y})^2} \quad (7.21)$$

ولما كانت $\Sigma \hat{q}_t = 0$ و $\Sigma (\hat{q}_t \hat{Y}_t) = 0$ فإننا نحصل على :

$$\begin{aligned} \sum (\hat{Y}_t u_t^*) &= \sum (\hat{Y}_t u_t) + b_1 \sum (\hat{Y}_t \hat{q}_t) = \sum (\hat{Y}_t u_t), \\ \sum (\hat{Y}_t u_t^*) &= \bar{Y} \sum u_t + \bar{Y} b_1 \sum \hat{q}_t = \bar{Y} \sum u_t. \end{aligned}$$

ويمكننا الآن أن نعبر عن \hat{b}_1 على النحو :

$$\hat{b}_1 = b_1 + \frac{\sum (\hat{Y}_t - \bar{Y}) u_t}{\sum (\hat{Y}_t - \bar{Y})^2} \quad (7.22)$$

وينتج عن ذلك، حدسياً، في الأقل، أن تباين العينة الكبيرة لـ \hat{b}_1 سيتضمن تباين u_t وليس u_t^* . ونترك للقارئ أن يثبت أنه يمكن، للأسباب نفسها، التعبير عن \hat{b}_0 بدلالة u_t .

وعلى الرغم من أن $\hat{\sigma}_u^2$ يكون، عادة، غير معلوم، فإنه يمكننا الحصول على مقدرات متسقة لتباينات \hat{b}_0 و \hat{b}_1 عن طريق إحلال المقدر $\hat{\sigma}_u^2$ المتسق محل σ_u^2 :

$$\hat{\sigma}_{\hat{b}_0}^2 = \hat{\sigma}_u^2 \left[\frac{\sum \hat{Y}_t}{n \sum (\hat{Y}_t - \bar{Y})^2} \right] \quad (7.23)$$

و

$$\hat{\sigma}_{\hat{b}_1}^2 = \hat{\sigma}_u^2 \left[\frac{1}{\sum (\hat{Y}_t - \bar{Y})^2} \right] \quad (7.24)$$

وأخيراً، يمكن اختبار الفرضيات أو بناء فترات الثقة عن طريق افتراض أن النسبة $(\hat{b}_i - b_i) / \hat{\sigma}_{\hat{b}_i}$ حيث $(i=0 \text{ و } 1)$ متغير طبيعي معياري. ومرة أخرى، نجد، عند التطبيق، أن العينات ذات حجم محدود، ولذا فإن مثل هذه النتائج هي نتائج تقريبية، والسبب وراء التعقيد في هذه الحالة هو عدم خطية كل من \hat{b}_1 و \hat{b}_0 بالنسبة للأخطاء العشوائية من خلال اعتمادها على \hat{Y}_1 (على سبيل المثال، انظر (7.22).

(٧-٣) نظم المعادلات : مناقشة أكثر عمومية*

تحديد النموذج

سنعمم الآن مناقشتنا لنماذج المعادلات الآتية. افترض نموذج الانحدار المتعدد الذي يتكون من المعادلات :

$$Y_{1t} = b_0 + b_1 Y_{2t} + b_2 Y_{3t} + a_1 X_{1t} + \dots + a_k X_{kt} + u_t, \quad (7.25)$$

$$Y_{2t} = c_0 + c_1 Y_{3t} + c_2 Y_{1t} + d_1 Z_{1t} + \dots + d_l Z_{lt} + \varepsilon_t, \quad (7.26)$$

$$Y_{3t} = g_0 + g_1 Y_{2t} + g_2 Y_{1t} + h_1 W_{1t} + \dots + h_s W_{st} + e_t, \quad (7.27)$$

حيث u_t, ε_t, e_t هي الأخطاء العشوائية. نفترض أن كل من هذه الأخطاء العشوائية له قيمة متوسطة صفرية :

$$E(u_t) = 0, \quad E(\varepsilon_t) = 0, \quad E(e_t) = 0,$$

وتباين ثابت :

$$E(u_t^2) = \sigma_u^2, \quad E(\varepsilon_t^2) = \sigma_\varepsilon^2, \quad E(e_t^2) = \sigma_e^2,$$

أي أنها غير مرتبطة ذاتياً كما أنها موزعة توزيعاً طبيعياً. إضافة إلى ذلك، سنفترض أن هذه الأخطاء العشوائية غير مرتبطة بكل من X'_t و Z'_t و W'_t الظاهرة في النظام (7.25) - (7.27). وعلى الرغم من أن الرموز المستخدمة في النموذج لاتدل

* لتبسيط العرض، نناقش نظم المعادلات في إطار نموذج ذي ثلاث معادلات، فقط. إلا أن النتائج تنطبق، أيضاً، على جميع النماذج ذات الاحجام المختلفة.

على ذلك، إلا أننا نسمح بإمكانية ظهور واحد أو أكثر من هذه المتغيرات في أكثر من معادلة واحدة، وبمعنى آخر فإن X_t ، Z_t و W_t قد تكون لها عناصر مشتركة. ينبغي أن يكون واضحاً عدم إمكانية افتراض أن الأخطاء العشوائية غير مرتبطة، أيضاً، بـ Y_{1t} ، Y_{2t} و Y_{3t} . وفي الحقيقة، وعموماً، فإن كل واحد من Y_t سوف يكون مرتبطاً بجميع الأخطاء العشوائية الثلاثة. على سبيل المثال يتضح لنا مباشرة من معادلة (7.25) أن Y_{1t} يعتمد على u_t ولذا، فإن هذين المتغيرين مرتبطان. فضلاً عن ذلك، تشير المعادلة (7.26) إلى أن Y_{2t} يعتمد على Y_{1t} ومن ثم على u_t . وهكذا تكون Y_{2t} ، عموماً، مرتبطة بـ u_t . وبالمثل، تتضمن (7.27) أن Y_{3t} ترتبط، عموماً، بـ u_t . ويمكن توسيع المناقشة هذه لتوضيح التغيرات غير الصفيرية بين Y_{1t} ، Y_{2t} و Y_{3t} مع الأخطاء العشوائية الأخرى ε_t ، e_t . وباختصار، إذا اشترك متغيران في علاقة ذات اتجاهين فإن كل متغير سوف يكون مرتبطاً بالخطأ العشوائي في معادلة المتغير الآخر.

في هذا النموذج ذي المعادلات الثلاث تسمى Y_{1t} ، Y_{2t} و Y_{3t} بالمتغيرات الداخلية endogenous variables، تحدد قيمتها في الزمن t بواسطة النموذج المحدد في المعادلات (7.25-7.27). وهي المتغيرات التي حاول نمودجنا تفسيرها. ويطلق على المتغيرات الأخرى X_t ، Z_t و W_t بالمتغيرات المحددة قديماً مسبقاً Predetermined variables. وهذه المتغيرات غير مرتبطة بالأخطاء العشوائية، كما أن قيمها في الفترة t لا تتحدد بواسطة النموذج في الفترة t . وسنوضح بعد قليل كيف تحدد قيمها. وفي الوقت نفسه ينبغي أن يكون واضحاً أن قيم هذه المتغيرات المحددة مسبقاً مع قيم الأخطاء العشوائية تحددان معاً قيم المتغيرات الداخلية في النموذج في الفترة t . فمثلاً، نجد أن مجموعة المعادلات الخطية مثل (7.25-7.27) يمكن حلها للحصول على قيم المتغيرات الداخلية Y_{1t} و Y_{2t} و Y_{3t} بدلالة المتغيرات المحددة قيمها مسبقاً والأخطاء العشوائية. وهكذا، يمكننا، عن طريق حل هذا النموذج، الحصول على Y_t في الشكل :

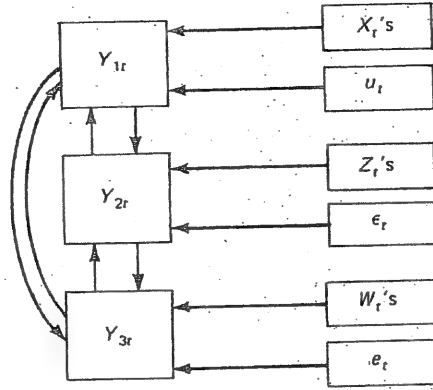
$$Y_{1t} = l_0 + \sum_{i=1}^k l_i X_{it} + \sum_{i=k+1}^{k+r} l_i Z_{it} + \sum_{i=k+r+1}^{k+r+s} l_i W_{it} + \alpha_1 u_t + \alpha_2 \varepsilon_t + \alpha_3 e_t, \quad (7.28)$$

$$Y_{2t} = m_0 + \sum_{i=1}^k m_i X_{it} + \sum_{i=k+1}^{k+r} m_i Z_{it} + \sum_{i=k+r+1}^{k+r+s} m_i W_{it} + \gamma_1 u_t + \gamma_2 \varepsilon_t + \gamma_3 e_t, \quad (7.29)$$

$$Y_{3t} = d_0 + \sum_{i=1}^k d_i X_{it} + \sum_{i=k+1}^{k+r} d_i Z_{it} + \sum_{i=k+r+1}^{k+r+s} d_i W_{it} + \beta_1 u_t + \beta_2 \varepsilon_t + \beta_3 e_t, \quad (7.30)$$

وهكذا يتضح لنا أن قيم المتغيرات المحددة قيمها مسبقا وقيم الأخطاء العشوائية تحدان، تماما، قيم المتغيرات الداخلية، ويتضمن ذلك أن للمتغيرات المحددة مسبقا وللأخطاء العشوائية سمات مشتركة، فكل منهما لا يتحدد بوساطة النموذج، كما أنهما ضروريان لتحديد قيم المتغيرات الداخلية، ولكن، هناك اختلاف رئيسي بينهما حيث إننا نفترض، على العكس من الأخطاء العشوائية، أننا نعرف أو نشاهد قيم المتغيرات المحددة مسبقا كل فترة زمنية. ويعني ذلك أنه إذا توافرت لدينا مشاهدات حول تقيم خطأ عشوائي معين فإنه يمكن اعتبار أن هذا الخطأ العشوائي محددة قيمته مسبقا. وبالمقابل، قد يمكننا اعتبار الخطأ العشوائي متغيرا محددا مسبقا غير مشاهد له متوسط صفري وأنه غير مرتبط بالمتغيرات المحددة مسبقا المشاهدة كافة.

وعلى سبيل الخلاصة، نعرض في الشكل (٧-٢) تمثيلا منظما لهيكل العلاقة السببية لمجموعة العلاقات في (7.25) - (7.27). نلاحظ (كما هو موضح بالأسهم) أن المتغيرات المحددة مسبقا والأخطاء العشوائية تؤثر مباشرة في قيم المتغيرات الداخلية ولكنها لا تتأثر، بدورها، بها. وعلى العكس، هناك علاقة متبادلة (أو اعتماد متبادل) بين المتغيرات الداخلية، حيث تعتمد Y_{1t} ، مثلا، على كل من Y_{2t} و Y_{3t} ، ولكنها تؤثر، أيضا، في كل منهما.



شكل رقم (٧-٢)

طبيعة المتغيرات المحددة مسبقا

تأخذ المتغيرات المحددة مسبقا أحد شكلين، الأول : المتغيرات الخارجية Exogenous variables، وكما أشرنا في مناقشتنا من قبل لنموذج الاستهلاك المبسط (7.1)-(7.2) فإن هذا المتغيرات تحدد قيمها بواسطة قوى خارجة عن النموذج، ويفترض أن قيم المتغيرات الخارجية تعتمد على متغيرات ليست مرتبطة بأي شكل بالمتغيرات الداخلية أو بالأخطاء العشوائية في النموذج. وبمعنى آخر فإن هذه المتغيرات خارج نطاق التحليل. * وببساطة، نأخذ قيمتها كما تعطي لنا بدون أية محاولة لتفسيرها. ومن الأمثلة لأحد المتغيرات التي يمكن أن تعد متغيرا خارجيا في معادلة الاستهلاك حجم الأسرة. هذا المتغير يمكن اعتباره تغذية من النظام الاجتماعي للنموذج الاقتصادي، ويعد هذا المتغير متغيرا مهما في تحديد متغيرات اقتصادية، مثل الاستهلاك ولكن (في الأقل، في الفترة القصيرة) قد لا يتأثر، بدوره، بالمتغيرات الاقتصادية في نموذجنا هذا. وأحد الأمثلة الأخرى للمتغيرات الخارجية قد يكون كمية الامطار السنوية في معادلة لتفسير إنتاج القمح. وفي

* تحديد نطاق التحليل ليست مهمة سهلة، والمشكلة هي أن التحليل الأكثر شمولاً يتطلب «مجالاً أوسع» ولكن هذا، بدوره، يزيد من تعقيد النموذج.

الحالة الأولى (حجم الأسرة)، قد نفكر في توسيع نطاق التحليل لتفسير المتغير الخارجي، فمثلا قد يعتمد حجم الأسرة على فرص العمل وهلم جرا. وإذا كان الأمر كذلك فإن هذا المتغير لم يعد متغيرا خارجيا. وفي الحالة الثانية، قد تعد كمية الأمطار السنوية خارج نطاق أي توسيع مقبول للتحليل.

والشكل الثاني من المتغيرات المحددة مسبقا هو المتغير الداخلي المبطل، ففي نموذج الاستهلاك، على سبيل المثال :

$$C_t = b_0 + b_1 Y_t + b_2 Y_{t-1} + u_t, \quad (7.31)$$

$$Y_t = C_t + I_t, \quad (7.32)$$

يمكننا أن نرى أنه، في ظل افتراضاتنا المعتادة، يكون Y_t مرتبطا بالخطأ العشوائي u_t لأننا وجدنا عند حل المعادلات للحصول على Y_t ، (كما لوحظ من قبل) أن Y_t تعتمد على u_t :

$$Y_t = \left(\frac{b_0}{1-b_1} \right) + \left(\frac{b_2}{1-b_1} \right) Y_{t-1} + \left(\frac{1}{1-b_1} \right) I_t + \left(\frac{1}{1-b_1} \right) u_t. \quad (7.33)$$

ولكن ذلك ليس صحيحا بالنسبة لـ Y_{t-1} لأن قيمة Y_{t-1} محددة من الفترة الزمنية السابقة، ولأن قيمتها لا يمكن أن تعتمد على C_t و u_t أو قيمة أي متغير آخر في الفترة t . على سبيل المثال، فإنه، عند ابدال t بـ $(t-1)$ في (7.33)، نجد أن Y_{t-1} تعتمد على u_{t-1} وليس u_t .

$$Y_{t-1} = \left(\frac{b_0}{1-b_1} \right) + \left(\frac{b_2}{1-b_1} \right) Y_{t-2} + \left(\frac{1}{1-b_1} \right) I_{t-1} + \left(\frac{1}{1-b_1} \right) u_{t-1}.$$

ويتضمن هذا أنه، إذا كانت قيمة الخطأ العشوائي u_t تحدد مستقلة في كل فترة زمنية (حتى يكون u_t غير مرتبط ذاتيا)، فإن Y_{t-1} و u_t يجب أن يكونا غير مرتبطين. وباختصار، فإنه، وبسبب أن قيمة Y_{t-1} ليست محددة بواسطة النموذج خلال الفترة t ولأنها غير مرتبطة بالخطأ العشوائي في الزمن t ، فإنه يمكننا أن

نصف Y_{t-1} على أنه متغير محدد مسبقاً* . وافترضنا الأساسي بالتغاير الصفري بين المتغيرات المحددة مسبقاً والأخطاء العشوائية صحيحاً لكل المتغيرات الخارجية والمتغيرات الداخلية المبطة في النموذج.

المعادلات الهيكلية ومعادلات الشكل المختزل

قبل أن نفكر في تعميم طريقتنا في التقدير، نحتاج لتوضيح أخير. تسمى المعادلات الأساسية مثل (7.25)، (7.26) و (7.27) والتي تفسر سلوك المتغيرات الداخلية والتشابك فيما بينها المعادلات الهيكلية Structural equations. هذه هي المعادلات التي تقترحها لنا النظرية الاقتصادية. مثال ذلك أننا في نموذجنا البسيط لتحديد الدخل عرضنا نظاماً من معادلتين هيكليتين :

$$C_t = b_0 + b_1 Y_t + u_t. \quad (7.1)$$

$$Y_t = C_t + I_t, \quad (7.2)$$

تتضمن هذه المعادلات الافتراضات النظرية بأن الإنفاق الاستهلاكي يعتمد على الدخل (7.1) وأن الناتج الكلي ينتقل إلى مستوى توازني يتعادل فيه مع الطلب الكلي (7.2). ومرة أخرى، فإن المعادلات الهيكلية هي تعبير اصطلاحي لنموذجنا الاقتصادي الأساسي.

إذا حلت المعادلات الهيكلية للحصول على المتغيرات الداخلية بدلالة المتغيرات المحددة مسبقاً والأخطاء العشوائية مثل الموجودة في (7.22) - (7.30) فإنه يطلق على المعادلات الناتجة المعادلات ذات الشكل المختزل Reduced form equations. وتصف

* والاستثناء من ذلك هو الحالة التي تكون فيها الأخطاء العشوائية مرتبطة ذاتياً، وفي مثل هذه الحال، يمكن أن يكون المتغير الداخلي المبطة مرتبطاً بالأخطاء العشوائية في الفترة t في النموذج. وللقارئ المهتم، يناقش ملحق هذا الفصل كيف يمكن تقدير النموذج الذي يعاني مشكلة النظم والأخطاء العشوائية المرتبطة ذاتياً في الوقت نفسه، وقد تجد ذلك مفيداً في التعرف على كيفية تعامل الاقتصاديين القياسيين مع أكثر من مشكلة تقدير واحدة في الوقت نفسه.

هذه المعادلات كيف يحدد المتغير الداخلي بوساطة كل من المتغيرات المحددة مسبقا والأخطاء العشوائية.

وعلى سبيل المثال، إذا قمنا بحل المعادلات الهيكلية (7.1) و (7.2) للحصول على المتغيرين الداخليين فإننا نحصل على :

$$C_t = \frac{b_0}{1-b_1} + \frac{b_1 I_t}{1-b_1} + \frac{u_t}{1-b_1}, \quad (7.34)$$

$$Y_t = \frac{b_0}{1-b_1} + \frac{I_t}{1-b_1} + \frac{u}{1-b_1}. \quad (7.35)$$

ويمكننا إعادة كتابة (7.34) و (7.35) في الشكل :

$$C_t = a_0 + a_1 I_t + q_t, \quad (7.34A)$$

$$Y_t = d_0 + d_1 I_t + q_t, \quad (7.35A)$$

حيث إن $a_0 = d_0 = b_0 / (1-b_1)$ ، $a_1 = b_1 (1-b_1)$ و $d_1 = 1 / (1-b_1)$ وأخيرا $q_t = u_t / (1-b_1)$ ويطلق على المعادلات (7.34A) و (7.35A) أو المعادلات (7.34) و (7.35) معاملات الشكل المختزل لكل من C_t و Y_t على الترتيب. لاحظ أن هذه المعادلات المختزلة يترتب عليها أن كلا من C_t و Y_t يتحددان تماما بوساطة I_t و u_t .

وعلى سبيل مثال آخر، افترض نموذج الأجور - الأسعار التالي :

$$\dot{W}_t = a_1 + b_1 \dot{P}_t + c_1 R_{t-1} + \varepsilon_{1t}, \quad (7.36)$$

$$\dot{P}_t = a_2 + b_2 \dot{W}_t + b_3 \dot{T}_{t-1} + \varepsilon_{2t}, \quad (7.37)$$

حيث إن \dot{W} و \dot{P} هي معدل التغير في الأجور النقدية والأسعار الاستهلاكية على الترتيب، R_{t-1} هو معدل البطالة في آخر فترة زمنية، \dot{T}_{t-1} هو معدل التغير في أسعار المواد الأولية في الفترة الزمنية و ε_{1t} ، ε_{2t} هي الأخطاء العشوائية. المعادلات (7.36) و (7.37) هي المعادلات الهيكلية بالنموذج، فهي تمثل الافتراضات بأن تغيرات الأجور تعتمد على تغيرات الأسعار، وعلى ظروف سوق العمل (كما هي ممثلة بمعدل البطالة في الفترة الزمنية السابقة) بينما تعتمد تغيرات الأسعار، بدورها،

على تغيرات الأجور والتغيرات في التكاليف الأخرى التي تمثلها بالتغيرات في أسعار المواد الأولية خلال الفترة الزمنية السابقة. وتكون المتغيرات الداخلية في النموذج هي \dot{W}_t و \dot{P}_t بينما تكون المتغيرات المحددة مسبقاً هي $R_{(t-1)}$ و $\dot{T}_{(t-1)}$. في هذه الحال نجد أن المتغيرات المحددة مسبقاً هي متغيرات خارجية مبطأة*، ولا توجد هناك متغيرات داخلية مبطأة، وأخيراً، إذا حللنا النموذج (7.36) و (7.37) للحصول على كل من \dot{W}_t و \dot{P}_t فإننا نحصل على معادلات الشكل المختزل :

$$\dot{W}_t = d_0 + d_1 \dot{T}_{t-1} + d_2 R_{t-1} + d_3 \varepsilon_{1t} + d_4 \varepsilon_{2t}, \quad (7.38)$$

$$\dot{P}_t = e_0 + e_1 \dot{T}_{t-1} + e_2 R_{t-1} + e_3 \varepsilon_{1t} + e_4 \varepsilon_{2t}, \quad (7.39)$$

حيث :

$$d_0 = \frac{a_1 + b_1 a_2}{1 - b_1 b_2}, \quad d_1 = \frac{b_1 b_3}{1 - b_1 b_2}, \quad d_2 = \frac{c_1}{1 - b_1 b_2}, \quad d_3 = \frac{1}{1 - b_1 b_2}, \quad d_4 = \frac{b_1}{1 - b_1 b_2},$$

$$e_0 = \frac{a_2 + b_2 a_1}{1 - b_1 b_2}, \quad e_1 = \frac{b_3}{1 - b_1 b_2}, \quad e_2 = \frac{b_2 c_1}{1 - b_1 b_2}, \quad e_3 = \frac{b_3}{1 - b_1 b_2}, \quad e_4 = \frac{1}{1 - b_1 b_2}.$$

لاحظ أن معادلات الشكل المختزل هي معادلات خطية في كل من المتغيرات المحددة مسبقاً وفي الأخطاء العشوائية.

(٧-٤) طريقة المربعات الصغرى ذات المرحلتين : تعميم

نظرة عامة

دعنا نعود الآن إلى مشكلتنا في التقدير. افترض أن لدينا مجموعة المعادلات (7.25)-(7.27) ونريد تقدير معاملات المعادلة (7.25) من هذا النموذج :

* يعترض بعضهم على «نطاق التحليل الضيق» في (7.36) و (7.37)، أي أن تحليلاً كاملاً للعلاقات بين الأجور - والأسعار ينبغي أن يتضمن معادلات تفسر معدل البطالة (بمعنى أن معدل البطالة لا ينبغي أن يعد متغيراً خارجياً في نموذج يفسر العلاقة بين الأجور والأسعار). ولكن، كما ذكرنا من قبل، فإن تكلفة مثل هذا النموذج الأكثر شمولاً ستكون زيادة في درجة تعقيد. ومرة أخرى، فإن تحديد الفاصل الذي يحدد نطاق التحليل يعتمد على تقدير الباحث.

$$Y_{1t} = b_0 + b_1 Y_{2t} + b_2 Y_{3t} + a_1 X_{1t} + \dots + a_k X_{kt} + u_t, \quad (7.25)$$

$$Y_{2t} = c_0 + c_1 Y_{3t} + c_2 Y_{1t} + d_1 Z_{1t} + \dots + d_r Z_{rt} + \varepsilon_t, \quad (7.26)$$

$$Y_{3t} = g_0 + g_1 Y_{2t} + g_2 Y_{1t} + h_1 W_{1t} + \dots + h_s W_{st} + e_t, \quad (7.27)$$

تتكون طريقة م ص م من الخطوات التالية : أولاً ، يجري انحدار للمتغيرات الداخلية في النموذج وهي Y_{2t} و Y_{3t} (أي المتغيرات المستقلة المرتبطة بالخطأ العشوائي) على كل المتغيرات المحددة مسبقاً في هذا النموذج المكون من المعادلات الثلاث . وبالتحديد ، فإننا سنقدر ، بوساطة طريقة التقدير المعتادة ، المعادلة :

$$Y_{2t} = m_0 + m_1 X_{1t} + \dots + m_k X_{kt} + m_{(k+1)} Z_{1t} + \dots + m_{(k+r)} Z_{rt} + m_{(k+r+1)} W_{1t} + \dots + m_{(k+r+s)} W_{st} + \theta_{1t}, \quad (7.40)$$

حيث θ_{1t} الخطأ العشوائي* ، وبالمثل ، سنجري انحداراً لـ Y_{3t} على المجموعة نفسها من المتغيرات المستقلة . ثم نستخدم بعد ذلك هذه المعادلات المقدرة لتحديد القيم المحسوبة \hat{Y}_{2t} و \hat{Y}_{3t} . ثانياً ، إحلال \hat{Y}_{2t} و \hat{Y}_{3t} محل Y_{2t} و Y_{3t} على الترتيب في المعادلة (7.25) . ونضع نجمة على الخطأ العشوائي لتمييزه ، وبعد ذلك ، نقوم بإيجاد المعادلات الطبيعية للمرحلة الثانية وحلها للحصول على مقدراتنا \hat{b}_0 ، \hat{b}_1 و \hat{b}_2 وايضاً \hat{a}_1 ، \hat{a}_2 ، ... و \hat{a}_k ، ومرة أخرى يتبين لنا أن \hat{Y}_{2t} و \hat{Y}_{3t} هي مقدراتنا للجزء من $Y_{2t} Y_{3t}$ ، على الترتيب ، غير المرتبطة بخطأ معادلتنا العشوائي ، u_t .

تأطير

لمعرفة ذلك بصورة نظرية أوضح ، نلاحظ أنه ، طالما أن نموذجنا ذا المعادلات الثلاث خطي فإن الحل (أو معادلة الشكل المختزل) لـ Y_{2t} ، مثلاً ، سوف يكون خطياً في كل من $X_1's$ و $Z_1's$ و $W_1's$ ، وأيضاً ، في u_t ، ε_t و e_t ، ولتبسيط هذه

* كما سنبين فيما بعد ، فإن الخطأ العشوائي θ_{1t} هو مجموع موزون للأخطاء العشوائية الثلاث $(\phi_1 u_t + \phi_2 \varepsilon_t + \phi_3 e_t)$ كما هو الحال في المعادلة (7.29).

الرموز، نرمز لمجموع كل الحدود في معادلة الشكل المختزل (انظر (7.25) باستثناء الأخطاء العشوائية كـ Y_{2t}^m أي أن Y_{2t}^m هو توليفة خطية من X_t ، Z_t و W_t .*

$$Y_{2t}^m = m_0 + m_1 X_{1t} + \dots + m_k X_{kt} + m_{(k+1)} Z_{1t} + \dots + m_{(k+r)} Z_{rt} + m_{(k+r+1)} W_{1t} + \dots + m_{(k+r+s)} W_{st} + W_{st}. \quad (7.41)$$

يمكننا الآن أن نعبر عن معادلة الشكل المختزل لـ Y_{2t} ببساطة أكثر على النحو:

$$Y_{2t} = Y_{2t}^m + \gamma_1 u_t + \gamma_2 \varepsilon_t + \gamma_3 e_t, \quad (7.42)$$

حيث إن $\gamma_1, \gamma_2, \gamma_3$ ثوابت.

لاحظ بعد ذلك أن

$$E(\gamma_1 u_t + \gamma_2 \varepsilon_t + \gamma_3 e_t) = \gamma_1 E(u_t) + \gamma_2 E(\varepsilon_t) + \gamma_3 E(e_t) = 0,$$

طالما أن جميع الأخطاء العشوائية لها متوسطات صفرية، ودمج كل كل الأخطاء العشوائية الثلاثة في خطأ عشوائي واحد.

$$\theta_{1t} = \gamma_1 u_t + \gamma_2 \varepsilon_t + \gamma_3 e_t. \quad (7.43)$$

عندئذ، يمكننا كتابة (7.42) في الشكل:

$$Y_{2t} = Y_{2t}^m + \theta_{1t}, \quad (7.44)$$

حيث إن $E(\theta_{2t}) = v$ وبالمثل يمكننا أن نثبت أن:

$$Y_{3t} = Y_{3t}^m + \theta_{2t}, \quad (7.45)$$

حيث إن $E(\theta_{2t}) = 0$ و Y_{3t}^m هو توليفة خطية من X_t ، Z_t و W_t (انظر (7.30). أخيراً، كما فعلنا بالنسبة للحال البسيطة لمتغير مستقل واحد، يمكننا استخدام كل من (7.44) و (7.45) لإعادة كتابة أولى معادلاتنا الهيكلية (7.25) على النحو:

* على سبيل المثال، بالنسبة للنموذج (7.36) و (7.37)

$$\dot{W}_t^m = d_0 + d_1 \dot{T}_{(t+1)} + d R_{(t+1)}$$

حيث إن d_t عرفت (7.38).

$$Y_{1t} = b_0 + b_1 Y_{2t}^m + b_2 Y_{3t}^m + a_1 X_{1t} + \dots + a_k X_{kt} + q_{1t} \quad (7.46)$$

$$q_{1t} = u_t + b_1 \theta_{1t} + b_2 \theta_{2t}, \quad \text{حيث}$$

وهكذا يكون $E(q_{1t}) = 0$.

ينبغي أن يتضح الآن التناظر مع الحال البسيطة. فإذا كان لدينا مشاهدات عن Y_{2t}^m و Y_{3t}^m ، عندها، يمكننا، ببساطة، تقدير (7.46) بالطريقة المعتادة. والسبب في ذلك هو أنه طالما Y_{2t}^m و Y_{3t}^m تعتمدان، فقط، على X_{1t} ، Z_{1t} ، و W_{1t} ، وطالما أن هذه المتغيرات المحددة مسبقا غير مرتبطة بالأخطاء العشوائية، لذا، فإن كلا من Y_{2t}^m و Y_{3t}^m ينبغي ألا تكون مرتبطة بدورها بالأخطاء العشوائية. وكما في الحال المبسطة، فإننا لانعرف قيم كل من Y_{2t}^m و Y_{3t}^m ، ومن ثم، ينبغي أن نقدرها في البداية (بوساطة \hat{Y}_{2t} و \hat{Y}_{3t}) قبل أن نواصل التحليل كالمعتاد.

ينبغي أن نشير هنا إلى أن المقدرات التي يحصل عليها باستخدام طريقة (م ص م)، كما في الحال المبسطة ذات المتغيرين، تتصف بالتحيز، عموما، إلا أنها متسقة. نلاحظ، أيضا، أنه، بوجود مقدرات متسقة، قد تتمكن من بناء مقدرات متسقة للخطأ العشوائي، مثلا في (7.25)، من الصيغة التالية:

$$\hat{u}_t = Y_{1t} - (\hat{b}_0 + \hat{b}_1 Y_{2t} + \hat{b}_2 Y_{3t} + \hat{a}_1 X_{1t} + \dots + \hat{a}_k X_{kt}). \quad (7.47)$$

بعد ذلك، قد نحصل على مقدر متسق لتباين الخطأ العشوائي بوساطة^{*}:

$$\hat{\sigma}_u^2 = \sum_{t=1}^n \frac{\hat{u}_t^2}{n-k-3}. \quad (7.48)$$

ندون هنا، بدون اثبات، أن إحلال المتغيرات الداخلية المحسوبة محل نظيراتها في النموذج يؤدي إلى صيغ لتباين مقدرات الملمات، كما لو أنها نتائج عينات كبيرة، تماما كما لو كان عندنا متغير مستقل واحد. فمثلا، بالرجوع إلى معادلة (7.25) يكون تباين العينة الكبيرة لمقدر \hat{b}_1 الناتج عن استخدام طريقة م ص م هو:

* نقسم على $(n-k-3)$ في (7.48) لأنه يوجد $(k+3)$ معلمات في نموذج الانحدار (7.46).

$$\sigma_{\hat{b}_1}^2 = \sigma_u^2 \left(\frac{1}{\sum \hat{v}_{1t}^2} \right), \quad (7.49)$$

حيث إن \hat{v}_{1t} بواقى انحدار \hat{Y}_{2t} على \hat{Y}_{3t} ، X_{1t}, \dots, X_{kt} . لاحظ مرة أخرى أن \hat{Y}_{2t} و \hat{Y}_{3t} هي متغيرات مستقلة في المرحلة الثانية بمعاملات في b_1, b_2 . وبالمثل نجد أن صيغة التباين في (7.49) تتضمن تباينات u_t وليس الخطأ العشوائي المناظر للمرحلة الثانية للانحدار $(u_t^* = u_t + b_1 \hat{\theta}_{1t} + b_2 \hat{\theta}_{2t})$. أخيراً، فطالما أن σ_u^2 غير معلومة، عادة، فإن المقدّر المتسق لتباين \hat{b}_1 يكون:

$$\sigma_{\hat{b}_1}^2 = \sigma_u^2 \left(\frac{1}{\sum \hat{v}_{1t}^2} \right), \quad (7.50)$$

وتوجد صيغ مشابهة لمقدرات المسمات الأخرى.

وباستخدام هذه الصيغ، يمكننا اختبار الفرضيات وبناء فترات الثقة عن طريق استخدام التوزيع الطبيعي كتقريب. وعملياً، ينبغي أن ينظر إلى مثل هذه النتائج على أنها نتائج تقريبية طالما أن حجم العينة محدود.

م ص م والمتغيرات المحذوفة

ينبغي علينا هنا توضيح أحد الأمور العملية. ففي (7.25)، تتكون المرحلة الأولى في ظل م ص م من انحدار كل متغير من المتغيرات المستقلة على جميع المتغيرات المحددة مسبقاً في النموذج. وقد توجد حالة لا تتوافر فيها البيانات عن جميع هذه المتغيرات. افترض، مثلاً، أنه ليست لدينا قيم مشاهدة عن Z_t و W_t من المعادلتين (7.26) و (7.27) على الترتيب. هذا لن يمنعنا، عموماً، من استخدام (م ص م) لتقدير (7.25)، حيث تكون المرحلة الأولى في هذه الحالة من انحدار Y_{3t} و Y_{2t} على المتغيرات المتاحة والمحددة مسبقاً كافة (أي تلك المتغيرات غير Z_t و W_t في مثالنا). ويمكننا، حيثئذ، أن نستخدم معادلات المرحلة الأولى للحصول على \hat{Y}_{2t}

و \hat{Y}_{3t} إذا قمنا باحلال \hat{Y}_{2t} و \hat{Y}_{3t} محل Y_{2t} و Y_{3t} على الترتيب، كما يمكننا أن نواصل التحليل للمرحلة الثانية بالطريقة السابقة نفسها. وتظل الصيغ والتائج كافة التي ناقشناها من قبل صحيحة. وهكذا، فقد يمكن استخدام طريقة م ص م مع بيانات أقل من كاملة.

سنقوم الآن بمناقشة أكثر تحليلية لتوضيح لماذا يحدث ذلك. دعنا نهتم أولاً بالمناقشة المتعلقة بـ Y_{2t} . لما كانت البيانات عن Z_t و W_t غير متوفرة، فإن انحدار المرحلة الأولى سيكون مشابها لـ (7.40) باستثناء غياب كل من Z_t و W_t .

$$Y_{2t} = \pi_0 + \pi_1 Y_{1t} + \dots + \pi_k X_{kt} + \pi_{(k+1)} Z_{2t} + \dots + \pi_{(k+r+1)} Z_{rt} + \pi_{(k+r)} W_{2t} + \dots + \pi_{(k+r+s-2)} W_{2t} + \theta_{3t}. \quad (7.51)$$

نوجد \hat{Y}_{2t} من مقدرات المعلومات لهذا النموذج التي حصل عليها عن طريق وضع الفروض التالية:

$$\sum \hat{\theta}_{3t} = 0, \quad \sum (\hat{\theta}_{3t} X_{jt}) = 0, \quad \sum (\hat{\theta}_{3t} Z_{it}) = 0, \quad \sum (\hat{\theta}_{3t} W_{mt}) = 0, \quad (7.52)$$

حيث $j = 1, 2, \dots, k$ ، $i = 2, \dots, r$ و $m = 2, \dots, s$ ، وبوضوح أكثر يكون لدينا:

$$\hat{Y}_{2t} = \hat{\pi}_0 + \hat{\pi}_1 Y_{1t} + \dots + \hat{\pi}_k X_{kt} + \hat{\pi}_{(k+1)} Z_{2t} + \dots + \hat{\pi}_{(k+r+1)} Z_{rt} + \hat{\pi}_{(k+r)} W_{2t} + \dots + \hat{\pi}_{(k+r+s-2)} W_{2t}. \quad (7.53)$$

وكالعادة، يمكن التعبير عن Y_{2t} على النحو:

$$Y_{2t} = \hat{Y}_{2t} + \hat{\theta}_{3t}, \quad (7.54)$$

حيث تعرف $\hat{\theta}_{3t}$ على أنها $(Y_{2t} - \hat{Y}_{2t})$. ومرة أخرى، ولأن \hat{Y}_{2t} تعتمد

على X_{1t}, \dots, X_{kt} و Z_{2t}, \dots, Z_{rt} و W_{2t}, \dots, W_{2t} ، فإنه ينتج عن (7.52) أن:

$$\sum (\hat{Y}_{2t} \hat{\theta}_{3t}) = 0. \quad (7.55)$$

يمكننا الحصول على صيغة مناظرة لـ \hat{Y}_{3t} عن طريق انحدار Y_{3t} على المجموعة نفسها من المتغيرات المحددة مسبقاً للحصول على Y_{3t} . مع ملاحظة أن:

$$Y_{3t} = \hat{Y}_{3t} + \hat{\theta}_{4t}, \quad (7.56)$$

حيث إن $\theta_{4t} = Y_{3t} - \hat{Y}_{3t}$ سيكون لدينا :

$$\sum (\hat{Y}_{3t} \hat{\theta}_{4t}) = 0. \quad (7.57)$$

وأخيرا فطالما أن \hat{Y}_{2t} و \hat{Y}_{3t} هما توليفان خطيان لـ X_{1t}, \dots, X_{kt} و Z_{1t}, \dots, Z_{2t} و W_{2t}, \dots, W_{st} ، فإنه يكون لدينا من (7.52) ومن الشروط المقابلة لـ $\hat{\theta}_{4t}$.

$$\sum (\hat{Y}_{2t} \hat{\theta}_{4t}) = 0 = \sum (\hat{Y}_{3t} \hat{\theta}_{4t}). \quad (7.58)$$

دعنا الآن نعود إلى المعادلة (7.25). بإحلال قيم Y_{3t} و Y_{2t} من معادلتني

(7.54)، (7.56) على الترتيب، نحصل على :

$$Y_{1t} = b_0 + b_1 \hat{Y}_{2t} + b_2 \hat{Y}_{3t} + a_1 X_{1t} + \dots + a_k X_{kt} + u'_t, \quad (7.59)$$

حيث إن :

$$u'_t = b_1 \hat{\theta}_{3t} + b_2 \hat{\theta}_{4t} + u_t.$$

ويستج عن (7.52) والشروط الخاصة بـ $\hat{\theta}_{4t}$ أن :

$$\sum u'_t = \sum u_t, \quad \sum (u'_t \hat{Y}_{2t}) = \sum (u_t \hat{Y}_{2t}), \quad (7.60)$$

$$\sum (u'_t \hat{Y}_{3t}) = \sum (u_t \hat{Y}_{3t}) \quad \sum (X_{jt} u'_t) = \sum (X_{jt} u_t),$$

حيث إن : $j = 1, \dots, k$. أي أن الشروط الموجودة في (7.60) تبين أنه، بهدف التقدير، قد يمكننا معالجة u'_t بالطريقة نفسها التي نعالج بها u_t .

على سبيل المثال، كما في حالة الانحدار البسيط، تبين الشروط (7.60) أنه،

لكي نقوم بتقدير (7.59)، فإنه ينبغي أن نضع الشروط التالية :

$$\sum \hat{u}'_t = \sum \hat{u}_t = 0 \quad \text{طالما أن } E(u_t) = 0$$

$$\sum (\hat{u}'_t \hat{Y}_{2t}) = \sum (\hat{u}_t \hat{Y}_{2t}) = 0 \quad \text{طالما أن } u_t \text{ غير مرتبطة بالمتغيرات المحددة} \quad (7.61)$$

$$\sum (\hat{u}'_t \hat{Y}_{3t}) = \sum (\hat{u}_t \hat{Y}_{3t}) = 0 \quad \hat{Y}_{3t} \text{ و } \hat{Y}_{2t} \text{ من مسبقا التي يعتمد عليها كل من}$$

$$\sum (\hat{u}_i' X_{ji}) = \sum (\hat{u}_i X_{ji}) = 0 \quad (j = 1, 2, \dots, k) \quad \text{لكل } E(u_i X_{ji}) = 0 \text{ أن } (طالما)$$

يمكن إثبات أنه، إذا قدرت معلمات (7.59) عن طريق فرض الشروط (7.61)، فإن طريقتنا المعتادة في التقدير ستؤدي إلى إيجاد مقدرات متسقة.

وباختصار وبمقارنة (7.59) بـ (7.25)، نجد أن كل ما ينبغي علينا عمله هو إحلال \hat{Y}_{2t} و \hat{Y}_{3t} محل Y_{2t} و Y_{3t} وتغيير رموز الأخطاء العشوائية ثم إكمال التحليل بالطريقة المعتادة.

وبالصدفة، فهذه السمة تعد من السمات الجيدة لطريقة م ص م. ولتقدير المعادلة موضع الاهتمام، فليس من الضروري القيام بتقدير نظام المعادلات بالكامل، كما أنه ليس ضروريا حتى تحديد ذلك النظام تحديدا كاملا أو الحصول على بيانات كاملة للمعادلات الأخرى في النموذج. وكل ما هو مطلوب هو وجود «مجموعة ملائمة» من المتغيرات المحددة مسبقا من نظام المعادلات يمكن استخدامها في انحدار المتغيرات المستقلة الداخلية (تلك المرتبطة بالخطأ العشوائي). وسنناقش فيما بعد ماذا نعني بالضبط «بالمجموعة الملائمة» من المتغيرات المحددة مسبقا. نشير هنا، فقط، إلى أن المجموعة الملائمة من المتغيرات المحددة مسبقا ينبغي أن تشمل كل المتغيرات المحددة مسبقا التي تظهر في المعادلة الجاري تقديرها. على سبيل المثال، في المعادلة (7.25)، نفترض أن Z_{1t} و W_{1t} ليستا متاحيتين لتحديد \hat{Y}_{2t} و \hat{Y}_{3t} . وحذفهما مقبول لأن أيًا من Z_{1t} و W_{1t} لا يظهر في المعادلة الهيكلية (7.25). ولكن إذا لم تستخدم X_{1t} في بناء \hat{Y}_{2t} و \hat{Y}_{3t} فإن الطريقة السابقة لن تؤدي إلى إيجاد مقدرات متسقة. وحتى نرى ذلك، نلاحظ أنه إذا لم تستخدم X_{1t} فلا يمكننا أن نستنتج من (7.52) أن $\Sigma(\hat{\theta}_{3t} X_{1t}) = 0$ لأنه لم يتم انحدار Y_{2t} على X_{1t} ، وفي الحقيقة، فإن هذا المجموع لن يكون صفرا، عموما. في هذه الحال، فإن المعادلة المناظرة لـ X_{1t} في (7.60) لن تكون صحيحة.

$$\sum (u_i' X_{1t}) \neq \sum (u_i X_{1t}).$$

ولن يكون مبررا وضع المجموع المقابل في (7.61) مساويا للصفر.

نلاحظ أخيراً أنه، على الرغم من عدم ضرورة استخدام جميع المتغيرات المحددة مسبقاً في النموذج في المرحلة الأولى من طريقة م ص م، إلا أن هناك فائدة من استخدامها جميعاً. وعموماً، يمكن إثبات أنه كلما زاد عدد المتغيرات المحددة مسبقاً المستخدمة في المرحلة الأولى من طريقة م ص م انخفضت تباينات العينة الكبيرة لمقدرات المعاملات التي يحصل عليها في المرحلة الثانية. وعلى الرغم من أن إثبات هذه النتيجة يقع خارج نطاق هذا الكتاب، فإن هذه النتيجة لا ينبغي أن تثير دهشتنا. إن استخدام «معلومات» إضافية في شكل متغيرات محددة مسبقاً في المرحلة الأولى ينبغي أن يؤدي إلى تحسين مقدرات المرحلة الثانية. نتجه الآن نحو مشكلة ما الذي يجعل مجموعة معينة من المتغيرات المحددة مسبقاً ملائمة؟ ويجب أن عن هذا السؤال من خلال مناقشة ما أصبح يعرف بمشكلة التمييز "the identification problem".

(٥-٧) مشكلة التمييز*

ربما نتذكر أننا أوضحنا، عند تقديم م ص م، أن إحدى مزايا هذه الطريقة هي أنه، ما دام في الإمكان تقدير المعادلة، فإن م ص م تعطي مقدرات متسقة للمعاملات. ولكن، قد تنشأ بعض الحالات المرتبطة بنظم المعادلات التي يكون من المستحيل عندها تقدير قيم بعض (أو ربما كل) المعلمات.

مثال (١)

افترض النموذج البسيط التالي لسوق إحدى السلع:

$$Q_t^d = a_0 + a_1 P_t + u_t, \quad (7.62)$$

$$Q_t^s = b_0 + b_1 P_t + \varepsilon_t, \quad (7.63)$$

* ترجع المساهمة الكلاسيكية في هذا الموضوع إلى: Franklin M. Fisher. *The Identification Problem In Econometrics*, (New-York, McGraw-Hill, 1966)، وتستند المناقشة في هذا البحث جزئياً إلى هذا العمل.

$$Q_t^d = Q_t^s = Q_t. \quad (7.64)$$

يتكون النموذج من ثلاث معادلات: معادلة الطلب (6.62) ومعادلة العرض (6.63) ومعادلة توازنية للسوق (6.64) حيث إن Q_t^d و Q_t^s هما الكمية المطلوبة والكمية المعروضة في الفترة t على الترتيب، P_t : سعر السلعة في الفترة t ، u_t و ε_t هما الخطآن العشوائيان المناظران، وكل منهما بمتوسط صفر وتباين ثابت، كما أن كلا منهما غير مرتبط ذاتيا وموزع توزيعا معتدلا. تسمح لنا المعادلة (6.69) بجعل كل من Q_t^d و Q_t^s معادلتان للكمية Q_t (وهي الكمية التي يتم تبادلها، فعلا، خلال الفترة t). ويتضمن ذلك أن السعر يتعدل، حتى تتساوى الكمية المطلوبة مع الكمية المعروضة خلال كل فترة زمنية. افترض أننا نرغب في تقدير المعلمات في معادلة الطلب. نلاحظ أولا أن P_t مرتبطة بـ u_t ، لذا، فإن طريقتنا المعتادة في التقدير سوف تعطينا مقدرات غير متسقة للمعاملات. ولئلين ذلك، اجعل الجانب الأيمن من المعادلة (7.62) يتساوى مع الجانب الأيمن من المعادلة (7.63) لنحصل على:

$$a_0 + a_1 P_t + u_t = b_0 + b_1 P_t + \varepsilon_t. \quad (7.65)$$

وبحل (7.65) للحصول على P_t ، نحصل على معادلة الشكل المختزل لـ P_t :

$$P_t = \frac{(b_0 - a_0)}{(a_1 - b_1)} + \frac{\varepsilon_t}{(a_1 - b_1)} - \frac{u_t}{(a_1 - b_1)}. \quad (7.66)$$

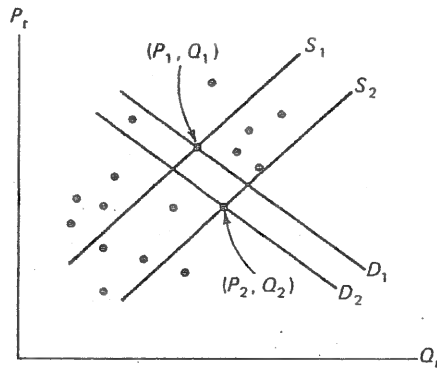
وبضرب (7.66) في u_t ، وأخذ القيمة المتوقعة نحصل على:

$$\begin{aligned} E(P_t u_t) &= E \left[\frac{(b_0 - a_0)u_t}{(a_1 - b_1)} + \frac{\varepsilon_t u_t}{(a_1 - b_1)} - \frac{u_t^2}{(a_1 - b_1)} \right] \\ &= \frac{(b_0 - a_0)}{(a_1 - b_1)} E(u_t) + \frac{E(\varepsilon_t u_t)}{(a_1 - b_1)} - \frac{E(u_t^2)}{(a_1 - b_1)} \\ &= 0 + \frac{\text{cov}(\varepsilon_t, u_t)}{(a_1 - b_1)} - \frac{\sigma_u^2}{(a_1 - b_1)}. \end{aligned} \quad (7.67)$$

طالما أنه لا يوجد سبب لتوقع أن تباين (ε_t, u_t) وتباين $(\sigma_u^2)u_t$ متساويان فإنه يمكننا أن نفترض، عموماً، أن:

$$E(P_t u_t) = \text{cov}(P_t, u_t) \neq 0. \quad (7.68)$$

وطالما أن (7.68) ليست، عموماً، مساوية للصفر، فإن ذلك يعني أن المتغير المستقل P_t مرتبط بـ u_t ، ولذا، تكون لدينا مشكلة نظم المعادلات الآتية. وحل المشكلة، افترض أننا نحاول تقدير (7.62) بواسطة م ص م. تتكون المرحلة الأولى، كما نذكر، من انحدار المتغير المستقل الداخلي (وهو هنا P_t) على جميع المتغيرات المحددة مسبقاً في النموذج. ولكن لمحة سريعة إلى نموذجنا ذي المعادلات الثلاث تبين عدم وجود متغيرات محددة مسبقاً*، ذلك أن معادلة P_t ذات الشكل المختزل (أي معادلة 7.66) لا تحتوي على متغيرات محددة سلفاً في جانبها الأيمن، ومن ثم، لا يمكننا هنا استخدام طريقة المربعات الصغرى ذات المرحلتين.

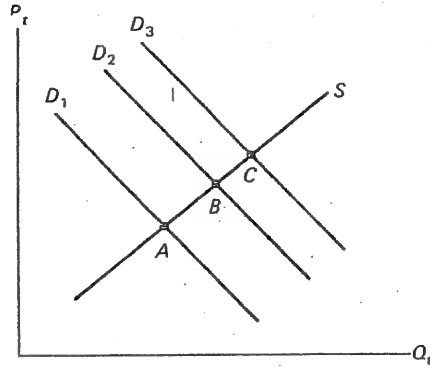


شكل رقم (٧-٣)

* في الحقيقة (وكما ستناقش فيما بعد)، يمكن اعتبار الحد الثابت متغيراً محدد مسبقاً. فإذا كان الأمر كذلك فإن معادلة P_t ذات الشكل المختزل (7.66) سوف تحتوي على متغير محدد مسبقاً. ولكن، (كما سنرى) فإننا سنظل غير قادرين على تقدير معادلة الطلب.

في هذه الحال، فإن معادلة الطلب (وأيضا، معادلة العرض) غير مميزة، بمعنى عدم إمكانية تقدير معالماتها. لاحظ أن مشكلة التمييز ليست مشكلة بيانات، لأنه مهما توافر لنا من بيانات حول السعر والكمية فإننا لن نكون قادرين على تقدير معاملات أي من معادلتى الطلب أو العرض. إن مشكلة التمييز هي مشكلة تحديد للنموذج، حيث إن هيكل النموذج وطبيعة المعلومات المتاحة هما اللذان يحولان دون إمكانية التقدير.

من المفيد هنا أن نعرض لمقولة بديهية ندعم بها صحة هذه النتيجة في الشكل (٣-٧)، يوجد لدينا شكل انتشار يحتوي على مشاهدات عن السعر والكمية عند نقاط زمنية مختلفة. وتمثل هذه النقاط المعلومات المتاحة التي ينبغي أن نستخدمها لتقدير منحنيات الطلب والعرض. والمشكلة التي نواجهها هي أن كل نقطة في الشكل تحدد بوساطة كلا المنحنيين الطلب والعرض. أي أن جداول الطلب والعرض تنتقل من فترة لأخرى (بسبب تغيرات ϵ_t و u_t) وينتج عن ذلك تغير في كل من السعر والكمية في السوق على مدى الزمن. ويتضح ذلك في الشكل (٣-٧) بوساطة مجموعتي منحنيات الطلب والعرض في الفترتين الزميتين الأولى والثانية اللتين تحددان معا Q_1, P_1 و Q_2, P_2 . وما يلاحظ هو مجموعة نقاط متشرة لتلك المجموعة في الشكل. ولكن لا توجد نقاط، في علمنا، نتجت عن تغير في الطلب، فقط. فإذا ماتوافرت مثل هذه النقاط فإنه يمكننا استخدام هذه النقاط لتقدير معاملات منحنى العرض. فعلى سبيل المثال، إذا علمنا في الشكل (٣-٧) أن النقاط A, B, C، قد أمكن الحصول عليها بوساطة ثلاثة منحنيات للطلب تتحرك على منحنى العرض نفسه، فإنه يمكننا، حدسيا، أن نستخدم هذه النقاط الثلاث لتقدير معالمات الميل والجزء الثابت من منحنى العرض. ولكن، في الشكل (٣-٧)، طالما أنه لا توجد لدينا طريقة يمكن بها أن نفصل بها تلك النقاط الناتجة عن تغير الطلب، فقط، فإننا غير قادرين على تقدير منحنى العرض. ولما كان المنطق نفسه يمكن أن يطبق على منحنى الطلب فلن تكون هناك طريقة في الشكل (٣-٧) يمكن أن «نميز» بها أي من منحنى الطلب أو العرض.



شكل (٧-٤)

مثال (٢)

افترض نموذج العرض - الطلب المعدل التالي :

$$Q_t^d = a_0 + a_1 P_t + a_2 P_{t-1} + u_t, \quad (7.69)$$

$$Q_t^s = b_0 + b_1 P_t + b_2 P_{t-1} + \varepsilon_t, \quad (7.70)$$

$$Q_t^d = Q_t^s = Q_t. \quad (7.71)$$

يشبه هذا النموذج (7.69) - (7.71) نموذجنا السابق للطلب والعرض باستثناء وحيد، وهو أن المتغير السعري المبطل يظهر في كل من معادلتَي الطلب والعرض. * وطالما أنه قد افترض أن الأخطاء العشوائية غير مرتبطة بـ P_{t-1} ، فإنها لن تكون مرتبطة كذلك بأي من u_t أو ε_t ، ولذا، قد يمكن اعتبارها متغيراً محدداً مسبقاً. ولكن متغير السعر الحالي، P_t ، يرتبط بكل من u_t أو ε_t ، ولذلك، ينبغي أن نستخدم طريقة معدلة لتقدير معادلات الطلب والعرض.

* قد يشير وجود P_{t-1} إلى أن سلوك المشتريين والبائعين يعتمد، تقريباً، على العادات أو المعلومات الماضية، أو بطريقة أخرى، فإن الأسعار المتوقعة في المستقبل تؤثر على القرارات التي تعتمد، بدورها، على تأثير كل من P_t و P_{t-1} في التوقعات.

افترض أننا نحاول استخدام طريقة م ص م لتقدير معلمات معادلة الطلب. ينبغي علينا أولاً أن نكون انحداراً لـ P_t على جميع المتغيرات المحددة مسبقاً في النموذج، في حالتنا هذه P_{t-1} ، ثم نوجد القيمة المحسوبة لـ P_t ، مثلاً:

$$\hat{P}_t = \hat{c}_0 + \hat{c}_1 P_{t-1}. \quad (7.72)$$

بعدئذ نحل \hat{P}_t محل P_t في (7.69) وفي المرحلة الثانية في تقدير المعادلة:

$$Q_t^d = a_0 + a_1 \hat{P}_t + a_2 P_{t-1} + u_t^*, \quad (7.73)$$

حيث تكون الكمية المطلوبة هي $Q_t^d = Q_t$. ولكن، في ضوء (7.72) يتعطل منهجنا في التقدير مرة أخرى، لأن قيم \hat{P}_t ستكون مرتبطة ارتباطاً خطياً تاماً مع P_{t-1} . هذه هي حالة الارتباط الخطي المتعدد التام التي لن تسمح لنا بإيجاد مقدرات منفصلة لكل من a_0 ، a_1 و a_2 . وتظل معادلتنا للطلب والعرض غير مميزتين.

مثال (٣)

دعنا نستكشف شكلاً ثالثاً لنموذجنا للعرض والطلب:

$$Q_t^d = a_0 + a_1 P_t + a_2 Y_t + u_t, \quad (7.74)$$

$$Q_t^s = b_0 + b_1 P_t + \varepsilon_t, \quad (7.75)$$

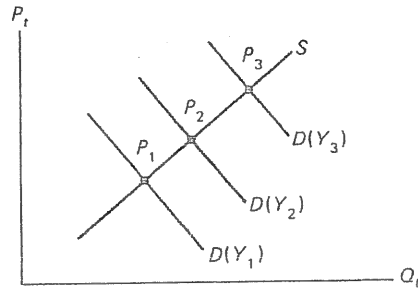
$$Q_t^d = Q_t^s = Q_t. \quad (7.76)$$

يشبه هذا النموذج نموذجنا الأصلي مع استثناء وحيد وهو إضافة متغير جديد يظهر في معادلة الطلب هو مستوى الدخل (Y_t). سنفترض، من أجل هذا المثال استقلال Y_t عن كل من u_t و ε_t . دعنا نرى إذا كان من الممكن في هذه الحالة تقدير أي من دالتي الطلب أو العرض. إذا اخذنا معادلة الطلب (7.74) أولاً، وجدنا أنه، بالإضافة إلى الحد الثابت، يوجد متغير واحد محدد مسبقاً في النموذج هو (Y_t). لذلك، نعمل انحداراً لـ P_t على Y_t للحصول على \hat{P}_t (مثلاً: $\hat{P}_t = \hat{h}_0 + \hat{h}_1 Y_t$) وبإحلال \hat{P}_t محل P_t في معادلة الطلب (7.74) نجد كما في الحالة السابقة أنه لا يمكننا الحصول على مقدرات لكل من a_0 ، a_1 و a_2 لأن لدينا مرة أخرى مشكلة الارتباط المتعدد الخطي

التام. إلا أنه يمكننا تقدير المعاملات في معادلة العرض (7.75). أي أنه إذا أحلت \hat{P}_t محل P_t في (7.75) فإنه يمكننا أن نستمر ونكمل طريقة م ص م، لأننا لانواجه مشكلة تعدد علاقات خطية. ونجد أنه يمكننا الحصول على مقدرات متسقة لـ b_0 و b_1 عن طريق إجراء انحدار لـ Q_t على \hat{P}_t . ويمكننا، في هذه الحال، تقدير معاملات معادلة العرض ولكن لا يمكننا ذلك بالنسبة لمعادلة الطلب.

يمكن أن يعطي تبرير أو توضيح بياني لهذه النتيجة. سيكون لدينا، مرة أخرى، شكل انتشار بين P_t و Q_t مشابه للشكل (٧-٣). ولكن، لدينا الآن معلومات إضافية مرتبطة بفصل هذه النقاط الناتجة عن تغير في الطلب فقط. ونرى، على وجه خاص، من تحديد معادلاتنا أن منحنيات الطلب والعرض تنتقل عندما تتغير الاخطاء العشوائية ε_t و u_t . ولكن، إضافة إلى هذا الأثر للأخطاء العشوائية، فإن منحنى الطلب، وليس منحنى العرض، سوف ينتقل إذا ماتغير Y_t . ويتضمن هذا، بالنسبة لشكل الانتشار، أنه إذا أمكن جعل الاخطاء العشوائية ثابتة عند الصفر، فإننا سنلاحظ مجموعة من النقاط تناظر القيم المختلفة لـ Y_t . وستمكنا هذه من تتبع منحنى العرض. وتظهر هذه الحالة في الشكل (٧-٥)، يمكننا في هذه الحالة أن نقدر منحنى العرض، إذا استطعنا الاعتماد على حقيقة أن الأخطاء العشوائية ليست مساوية للصفر، ولكنها تأخذ قيما مختلفة من فترة لأخرى. هذا التغير في الخطأ العشوائي قد أخذ في الحسبان، حدسيا، بواسطة افتراضنا بأن الاخطاء العشوائية مستقلة، وبالتالي غير مرتبطة بمستوى الدخل. ويمكننا هذا الشرط وافترض أن الاخطاء العشوائية لها وسط حسابي مساو للصفر من توزيع "average out" تأثير تلك الأخطاء العشوائية عن طريق استخدام طرقنا المعتادة في التقدير. وبالمقابل، فإن المنحنيات في الشكل (٧-٥) يمكن النظر إليها على أنها منحنيات متوسطة تناظر كل مستوى من مستويات الدخل.

أي أنه بالنسبة لأي قيمة من قيم Y_t تكون $E(u_t) = E(\varepsilon_t) = 0$ ، لذلك، فإن التأثير المتوسط للأخطاء العشوائية على موقع المنحنيات في الشكل (٧-٥) تكون صفرا. لاحظ، مع ذلك، عدم وجود مناقشة مماثلة تفيد إمكانية تقدير منحنى الطلب



شكل رقم (٧-٥)

لعدم وجود متغير في معادلة العرض يمكن (استنادا إلى عدم ظهوره في معادلة الطلب) أن يوجد انتقالات في منحنى العرض على مدى منحنى الطلب. في هذه الحالة الثالثة تكون معادلة العرض مميزة ولكن معادلة الطلب غير مميزة.

عرض أكثر عمومية

ستعرض الآن لبعض الأمثلة الأكثر عمومية. افترض أن المعادلة الهيكلية الأولى في نظام من المعادلات الآتية هي:

$$Y_{1t} = b_0 + b_1 X_{1t} + b_2 Y_{2t} + b_3 Y_{3t} + b_4 Y_{4t} + \varepsilon_t, \quad (7.77)$$

حيث Y_{1t} ، Y_{2t} و Y_{3t} متغيرات داخلية، X_{1t} متغير محدد مسبقا، وأن الخطأ العشوائي يحقق الافتراضات التقليدية كافة. افترض، أيضا، أن نظام المعادلات الذي تنتمي إليه هذه المعادلة يحتوي على متغير إضافي واحد محدد مسبقا، مثلاً، X_{2t} . في هذه الحالة لتقدير (7.77) سوف نستخدم منهج م ص م وذلك طالما أنه يتوقع أن تكون كل من Y_{2t} و Y_{3t} مرتبطة بـ ε_t بانحدار Y_{2t} و Y_{3t} على X_{1t} و X_{2t} نحصل على \hat{Y}_{2t} و \hat{Y}_{3t} .

$$\hat{Y}_{2t} = \hat{\gamma}_0 + \hat{\gamma}_1 X_{1t} + \hat{\gamma}_2 X_{2t}, \quad (7.78)$$

$$\hat{Y}_{3t} = \hat{\alpha}_0 + \hat{\alpha}_1 X_{1t} + \hat{\alpha}_2 X_{2t}. \quad (7.79)$$

فإذا عوضنا من (7.78) و (7.79) في المعادلة (7.77) فإن نموذج الانحدار للمرحلة الثانية يأخذ الشكل:

$$Y_{1t} = b_0 + b_1 X_{1t} + b_2 \hat{Y}_{2t} + b_3 \hat{Y}_{3t} + \varepsilon_t^*. \quad (7.80)$$

حيث ε_t^m هو الخطأ العشوائي الجديد لنموذجنا. سنحاول الآن أن نقدر (7.80) بالطريقة المعتادة. ولكن طريقتنا في التقدير ستفشل مرة ثانية لوجود الارتباط الخطي المتعدد التام بين المتغيرات المستقلة. ولتوضيح ذلك، نضرب المعادلة (7.78) بـ $\hat{\alpha}_2$ والمعادلة (7.79) بـ $\hat{\gamma}_2$ ثم نطرح الأخيرة من السابقة لها للتخلص من X_{2t} والحصول على:

$$\hat{Y}_{2t} \hat{\alpha}_2 = \hat{Y}_{3t} \hat{\gamma}_2 = (\hat{\alpha}_2 \hat{\gamma}_0 - \hat{\alpha}_0 \hat{\gamma}_2) + (\hat{\gamma}_1 \hat{\alpha}_2 - \hat{\alpha}_1 \hat{\gamma}_2) X_{1t}. \quad (7.81)$$

حيث تشير المعادلة (7.81) إلى وجود علاقة خطية تامة بين قيم \hat{Y}_{2t} ، \hat{Y}_{3t} و X_{1t} . وينتج عن ذلك أن إحدى معادلاتنا الطبيعية لن تكون مستقلة خطياً عن المعادلات الأخرى ومن ثم، لن يكون من الممكن تقدير معالم (7.80).

يمكننا أن نتفهم هذه النتيجة من خلال تفحص المعادلات الطبيعية المناظرة (7.80). نحصل على المعادلتين الطبيعيتين الأوليين عن طريق وضع $\Sigma \hat{\varepsilon}_t^* = 0$ و $\Sigma \hat{Y}_{3t} \hat{\varepsilon}_t^* = 0$. وتكون المعادلة الطبيعية الثالثة (والتي نحصل عليها عن طريق وضع $\Sigma (\hat{\varepsilon}_t^* \hat{Y}_{3t}) = 0$) مستقلة عن المعادلتين الأوليين لأن Y_{2t} تعتمد على X_{2t} [انظر إلى (7.78)]. فإذا لم تكن هذه المعادلة الأخيرة مستقلة عن المعادلتين السابقتين فإنها ستكون مساوية لـ $\hat{\gamma}_0$ مضروبة في المعادلة الأولى مضافاً إليها $\hat{\gamma}_1$ مضروبة في الثانية. بمعنى آخر فإن X_{2t} «تفصل» المعادلة الطبيعية الثالثة عن المعادلتين الأوليين. ولكن المشكلة تتعقد عندما نتجه إلى المعادلة الطبيعية الرابعة التي نحصل عليها عن طريق وضع $\Sigma \hat{Y}_{3t} \hat{\varepsilon}_t^* = 0$. فكوننا استخدمنا X_{2t} لعزل المعادلة الثالثة عن المعادلتين الأوليين. فهل من الممكن استخدامها مرة أخرى لفصل المعادلة الرابعة عن الثلاثة الأولى، وكما رأينا، فإن الإجابة هي بالنفي، لأن قيم \hat{Y}_{3t} مرتبطة ارتباطاً خطياً تاماً

بقيم X_{1t} و \hat{Y}_{2t} * ويمكن أن نتخيل أن كل معادلة خطية طبيعية مستقلة تتطلب معلومات جديدة - في صورة متغير جديد !! وطالما أنه لم يبق أي منها فإن معلومات (7.80) أو (7.77) ليست مميزة بسبب الحاجة إلى معادلة طبيعية رابعة مستقلة.**

افترض أن نظام المعادلات الذي يحتوي على (7.77) يشتمل على متغيرين إضافيين محددين مسبقاً (X_{2t} و X_{3t} مثلاً). حيث أن المرحلة الأولى سوف تشتمل (7.78) و (7.79) على متغير مستقل إضافي X_{3t} . فإذا حذفنا، كما سبق، X_{2t} من معادلات المرحلة الأولى فسوف يكون لدينا معادلة خطية تربط بين كل من \hat{Y}_{2t} و \hat{Y}_{3t} ، X_{3t} و X_{1t} . وبسبب هذا الحد X_{3t} ، لن تكون \hat{Y}_{3t} توليفة خطية تامة من X_{1t} و \hat{Y}_{2t} . في هذه الحال، لن يكون لدينا تعدد علاقات خطية تامة في المرحلة الثانية لطريقتنا م ص م. وستصبح المعادلة (7.77) مميزة (معرفة) الآن. ومن الواضح أن هذه النتيجة سوف تظل صحيحة إذا كان هناك أكثر من متغيرين إضافيين محددين مسبقاً يظهران في نظام المعادلات الذي تكون معادلة (7.77) إحدى معادلاته. أي أن معادلتنا موضع السؤال ستكون مميزة إذا كان عدد المتغيرات الإضافية المحددة مسبقاً [أو تلك التي لا تظهر في المعادلة (7.77)] أكبر من اثنين أو تساوي اثنين، وهو عدد المتغيرات المستقلة الداخلية.

* ينبغي أن تكون قادراً على استخدام (7.81) في إثبات أن المعادلة الطبيعية الرابعة سوف تساوي $(\hat{\alpha}_0\hat{Y}_2 - \hat{\alpha}_2\hat{Y}_0)/\hat{Y}_2$ مضروبة في الأولى، مضافاً إليها $(\hat{\alpha}_1\hat{Y}_2 - \hat{\alpha}_2\hat{Y}_1)/\hat{Y}_2$ مضروبة في الثانية مضافاً إليها $\hat{\alpha}_2/\hat{Y}_2$ مضروبة في الثالثة.

** مثال أكثر سهولة لهذه الحالة في النموذج:

$$Y_{1t} = b_0 + b_1Y_{2t} + b_3Y_{3t} + u_t,$$

حيث إن Y_{2t} و Y_{3t} هي متغيرات داخلية، ويوجد متغير واحد محدد مسبقاً، فقط، (مثلاً) في بقية نظام المعادلات. في هذه الحال، ستؤدي انحداراتنا للمرحلة الأولى والحسابات إلى إيجاد قيم لـ \hat{Y}_{2t} و \hat{Y}_{3t} . ولكي قيم كل من \hat{Y}_{2t} و \hat{Y}_{3t} ستكون مركباتها خطية تامة لـ، وستكون مترابطة ارتباطاً كاملاً مع بعضها البعض. ومن الواضح أن طريقتنا في التقدير ستفشل بسبب الارتباط الخطي المتعدد التام بين المتغيرات المستقلة.

نقطة أخيرة ينبغي أن نذكرها وهي أن المعادلة الطبيعية الأولى التي حصلنا عليها عن طريق وضع $\sum \varepsilon_i^* = 0$ يمكن أن نتخيل أنها تناظر الحد الثابت.* وبالمقابل، يمكننا أن نفكر في الحد الثابت كمتغير محدد مسبقا. وهكذا، نجد في مثالنا التالي أنه، إذا كانت المعادلة التي نقوم بقدرها لا تحتوي على حد ثابت، ولكن واحدة أو أكثر من المعادلات الأخرى في النموذج تحتوي عليه، فيمكننا أن نعد الحد الثابت واحدا من المتغيرات المحددة مسبقا في نظام المعادلات التي لا تظهر في معادلتنا التي نقدرها.

افترض، على سبيل المثال، النموذج التالي ذا المعادلتين:

$$Y_{1t} = b_1 Y_{2t} + b_2 X_t + \varepsilon_{1t}, \quad (7.82)$$

$$Y_{2t} = \alpha_1 + \alpha_2 Y_{1t} + \varepsilon_{2t}, \quad (7.83)$$

حيث إن X_t متغير محدد سلفا وإن ε_{1t} و ε_{2t} أخطاء عشوائية تحقق افتراضاتنا الأساسية كافة. عند اللمحة الأولى للنموذج، يشك القارئ أننا لن نكون قادرين على تطبيق م ص م على المعادلة (7.82)، طالما أنه يتوافر لدينا متغير واحد محدد مسبقا X_t في نظام المعادلات وأن ذلك المتغير يظهر في المعادلة التي نريد تقديرها. ولكننا نجد أنه، بسبب استبعاد «الحد الثابت» من المعادلة (7.82)، يمكننا تطبيق م ص م لتقدير (7.82) لأن الحد الثابت المستبعد يزودنا بالمتغير الإضافي المحدد سلفا والذي نحتاج إليه. فيمكننا، مثلا، من المرحلة الأولى لمنهج م ص م أن نحصل على:

$$\hat{Y}_{2t} = \hat{c}_1 + \hat{c}_2 X_t. \quad (7.84)$$

يتضمن الشرط (7.84) أن هناك ثلاث معادلات طبيعية يمكن الحصول عليها بواسطة وضع الفروض التالية

$$N_1: \sum \hat{\varepsilon}_{1t}^* = 0, \quad N_2: \sum (\hat{\varepsilon}_{1t}^* \hat{Y}_{2t}) = 0, \quad N_3: \sum (\hat{\varepsilon}_{1t}^* X_t) = 0, \quad (7.85)$$

* فمثلا، يمكن التعبير عن المعادلة (7.80) على النحو التالي:

$$Y_{1t} = b_0 X_{0t} + b_1 X_{1t} + b_2 \hat{Y}_{2t} + b_3 \hat{Y}_{3t} + \varepsilon_t^*,$$

حيث إن X_{0t} تساوي الواحد الصحيح في كل فترة زمنية ($X_{0t}=1$) لكل t . وهكذا، فإن الشرط $\sum \hat{\varepsilon}_t^* = 0$

يمكن أن نفكر بأنه $\sum (\hat{\varepsilon}_t^* X_{0t}) = 0$.

معادلتان، فقط، من هذه المعادلات الثلاث ستكونان مستقلتين، ويمكن طالما أن (7.82) لا تحتوي على الحدث الثابت أن نحتاج، فقط، إلى معادلتين طبيعيتين، وسوف نأخذ، ببساطة، المعادلات المتناظرة لكل من N_2 و N_3 في (7.85) لأن هذه الشروط تناظر المتغيرات المستقلة في المعادلة (7.82).

بيان عام*

استعانة بما ذكرناه في المبحث السابق دعنا نحاول الآن الوصول إلى قاعدة عامة للحصول على مقدرات متسقة باستخدام م ص م: نلاحظ أنه لكي تستخدم م ص م للحصول على مقدرات متسقة للمعاملات في معادلة الانحدار فلا بد أن لا يتجاوز عدد المتغيرات الداخلية المستقلة في المعادلة المراد تقديرها عدد المتغيرات المحددة مسبقا التي تظهر في النموذج ككل ولا تظهر في المعادلة المراد تقديرها. وبالمقابل، وعموما، تكون معادلة معينة في النموذج مميزة (ويمكن تقديرها باتساق) إذا كانت $K_2 \geq K_1$ ، حيث K_2 هي عدد المتغيرات المحددة مسبقا في النموذج والمستبعدة من المعادلة المعطاة، K_1 هي عدد المتغيرات الداخلية التي تظهر باعتبارها متغيرات مستقلة في تلك المعادلة.**

بالعودة إلى الأمثلة السابقة، نجد أنه في نموذجنا المبسط لتحديد الدخل، نستطيع أن نقدر معادلة الاستهلاك لأنه يوجد لدينا متغير داخلي واحد بوصفه متغيرا مستقلا Y_t ومتغير واحد محدد مسبقا I_t يوجد في النموذج ولا يظهر في

* تظل هذه القاعدة العامة المعطاة صحيحة في مجال تحليلنا. على سبيل المثال، لم نهتم بالحالات التي يكون فيها الباحث عالما بقيم تباينات معينة وعلاقات بين معاملات محددة. إلخ. ولكن من الإنصاف القول إن التحليل الذي عرضناه هنا هو الذي نواجهه في التطبيق في أغلب الحالات.

** ينبغي علينا أن نستخدم التعبير «عموما» لأن هذه الشروط في الواقع، شروط ضرورية، ولكنها غير كافية لتمييز المعادلة، انظر: (Arthur S. Goldberger, *Econometric Theory* (New York, Wiley, 1964), pp. 306-329. ومناقشة هذه الشروط الكافية خارج نطاق هذا الكتاب. ولكن، من الناحية التطبيقية، فإن الشروط التي ناقشناها هنا هي الشروط التي نهتم بها عامة عند التطبيق.

معادلة الاستهلاك. ولم نكن قادرين على تقدير أي من معادلتى الطلب أو العرض في النموذجين الأول والثاني من نماذج الطلب والعرض لأنه، في هاتين الحالتين، كان لدينا متغير مستقل واحد ولكن لم توجد متغيرات محددة مسبقا مستبعدة. وأخيرا، في نموذجنا الثالث للطلب والعرض، كانت معادلة العرض مميزة بحكم أن المتغير الداخلي المستقل الواحد في معادلة العرض P_t قد قابله متغير محدد مسبقا Y_t يظهر، فقط، في معادلة الطلب. ولكن، لا توجد متغيرات محددة مسبقا مستبعدة من دالة الطلب، ولذلك وجدنا أنفسنا غير قادرين على تقدير تلك المعادلة.

ومنطق هذه القاعدة ينبغي أن يكون واضحا من الأمثلة السابقة التي تعرضنا لها. فحينما لا يوجد عدد كاف من المتغيرات المحددة مسبقا والمستبعدة من المعادلة تتوقف طريقة م ص م للتقدير بسبب الارتباط الخطي المتعدد التام في المرحلة الثانية. دعنا نؤكد في الخلاصة أن هذه لا تمثل نقصا أو عيبا في طريقة م ص م على وجه التحديد، فطرق التقدير الأخرى، أيضا، غير قادرة على تزويدنا بمقدرات متسقة للمعادلات غير المميزة. إنه هيكل النموذج ذاته وطبيعة المعلومات المتاحة التي تمنع تفكيك التداخلات بين المتغيرات.

(٦-٧) تقدير م ص م: مثالان

حتى نعتاد على استخدام طريقة م ص م، نختم هذا الفصل بعرض مثالين، يتضمن المثال الأول منهما تقدير منحى طلب افتراضي على سلعة معينة. وسوف يسمح لنا هذا المثال بتتبع خطوات هذه الطريقة، بينما حصلنا على المثال الثاني من دراسة تطبيقية فعلية من إحدى الدوريات الاقتصادية الحديثة حول المالية العامة للمجالس المحلية.

نموذج للطلب والعرض

افترض أنه يوجد لدينا نموذج مبسط لسوق إحدى السلع الزراعية (الخرشوف،

مثلا):

$$Q_t^d = b_0 + b_1 P_t + b_2 Y_t + u_t, \quad (7.86)$$

$$Q_t^s = a_0 + a_1 P_{t-1} + a_2 W_t + \varepsilon_t, \quad (7.87)$$

$$Q_t^d = Q_t^s = Q_t. \quad (7.88)$$

تشير المعادلة (7.86)، معادلة الطلب إلى أن الكمية المطلوبة في الفترة t تعتمد على السعر P_t والدخل Y_t . بينما تبين معادلة العرض (7.87) أن إنتاج الخرشوف يعتمد على P_{t-1} ، سعر الفترة الزمنية السابقة، أي الفترة الذي يتخذ فيها القرار بالزراعة، وعلى الطقس W_t . سنستخدم الأمطار مقاسة بالبوصات مقياساً لأحوال الطقس. أخيراً، تتضمن المعادلة (7.88) وهي المعادلة التوازنية لسوق الخرشوف، أن السعر في الزمن t يتعدل ليساوي الكمية المطلوبة بالكمية المعروضة.

دعنا نفترض أن كلا من Y_t و W_t تتحدد بوساطة قوى خارج نطاق النموذج أي أنها متغيرات خارجية. افترض، أيضاً، أن الأخطاء العشوائية لها متوسطات صفرية، وتباينات ثابتة، وأنها موزعة توزيعاً طبيعياً، ولا تعاني من الارتباط الذاتي وأنها، أخيراً، مستقلة عن المتغيرات المحددة مسبقاً. افترض أننا نرغب في تقدير المعادلات في معادلة الطلب على الخرشوف. نلاحظ أولاً أن واحداً من المتغيرات المستقلة، P_t ، هو متغير داخلي يتوقع أن يكون مرتبطاً بالأخطاء العشوائية في النموذج. ولما كان استخدام طريقة المربعات الصغرى المعتادة سينجم عنها مقدرات غير متسقة للمعاملات، فإننا سنستخدم م ص م لتقدير (7.88). نلاحظ ثانياً أن معادلة الطلب مميزة، أي أنها تحتوي على متغير داخلي مستقل واحد، ولكن يوجد متغيران محددان مسبقاً P_{t-1} و W_t يظهران في النموذج ولا يظهران في معادلة الطلب نفسها. وبافتراض توافر بيانات حول هذه المتغيرات، فسيكون بالإمكان تقدير دالة الطلب باستخدام م ص م.

وقد جمعنا مجموعة من البيانات الافتراضية لمتغيرات النموذج التي تظهر في الجدول (٧-١). لاحظ وجود عشر مشاهدات، فقط، تكاد تكفي لتبرير استخدام منهج صحيح للحصول على نتائج للعينات الكبيرة. ونؤكد هنا على أن هدف هذا المثال، ببساطة، هو توضيح خطوات طريقة تقدير م ص م مع بعض الأرقام الفعلية.

أولى عملياتنا هي إجراء انحدار للمتغير المستقل في معادلة الطلب (أي P_t) على جميع المتغيرات المحددة مسبقاً في النموذج. فبالإضافة إلى الحد الثابت، توجد ثلاث متغيرات محددة مسبقاً: اثنان منها من المتغيرات الخارجية (Y_t, W_t) ومتغير داخلي مبطاً واحد (P_{t-1}). ولذا، فإن معادلة الانحدار المقدرة هي:

$$\hat{P}_t = -8.60 + 3.75 Y_t - 0.22 W_t + 0.42 P_{t-1} \quad (7.89)$$

نستخدم الآن (7.89) لحساب مجموعة «مصححة» من القيم لـ P_t ، أي أننا نعوض عن قيم Y_t, W_t, P_{t-1} في المعادلة (7.89) ونحسب القيمة المناظرة لـ P_t في كل فترة زمنية ($t = 1, \dots, 10$). تظهر هذه السلسلة من القيم المحسوبة لمتغير السعر (مع القيم المشاهدة P_t) في الجدول (٧-٢)*. أما المرحلة الثانية من طريقة م ص م فتتمثل بإحلال قيم \hat{P}_t محل قيم P_t في معادلة الطلب ومن ثم تقدير المعادلة الجديدة بالطريقة العادية. وبعمل انحدار Q_t على \hat{P}_t, Y_t نحصل على:

$$\hat{Q}_t = -39.9 - 1.3 \hat{P}_t + 9.5 Y_t, \quad (7.90)$$

(2.9) (3.7) (4.1)

حيث تكون الأرقام التي بين الأقواس هي القيم المطلقة المناظرة لنسب t . وهكذا يكون لدينا $\hat{b}_0 = 39.9$ ، $\hat{b}_1 = -1.3$ و $\hat{b}_2 = 9.5$.

فإذا كنا قد استخدمنا طريقة المربعات الصغرى المعتادة لتقدير معادلة الطلب على الخرشوف (أي استخدمنا P_t بدلاً من \hat{P}_t) فسوف نحصل على النتائج التالية:

$$\hat{Q}_t = -25.1 - 0.7 \hat{P}_t + 6.2 Y_t, \quad (7.91)$$

(1.9) (2.1) (2.8)

* لاحظ أنه يتوافر لدينا تسع مشاهدات، فقط، لـ \hat{P}_t ، حيث إننا فقدنا مشاهدة واحدة عند استخدام المتغير المبطاً \hat{P}_t في المعادلة (7.89).

جدول رقم (٧-١). بيانات افتراضية عن سوق الخرشوف*

الفترة (t)	الكمية (Q)	سعر الوحدة (P)	الدخل (Y)	الأمطار (W)
١	١١	٢٠	٨, ١	٤٢
٢	١٦	١٨	٨, ٤	٥٨
٣	١١	٢٢	٨, ٥	٣٥
٤	١٤	٢١	٨, ٥	٤٦
٥	١٣	٢٧	٨, ٨	٤١
٦	١٧	٢٦	٩, ٠	٥٦
٧	١٤	٢٥	٨, ٩	٤٨
٨	١٥	٢٧	٩, ٤	٥٠
٩	١٢	٣٠	٩, ٥	٣٩
١٠	١٨	٢٨	٩, ٩	٥٢

جدول رقم (٧-٢). الأسعار المقدرة والأسعار الفعلية

الفترة	P_t	\hat{P}_t
١	٢٠, ٠	-
٢	١٨, ٠	١٨, ٥
٣	٢٢, ٠	٢٣, ١
٤	٢١, ٠	٢٢, ٤
٥	٢٧, ٠	٢٤, ٢
٦	٢٦, ٠	٢٤, ٢
٧	٢٥, ٠	٢٥, ١
٨	٢٧, ٠	٢٦, ١
٩	٣٠, ٠	٢٩, ٨
١٠	٢٨, ٠	٢٩, ٧

* وحدات هذه المتغيرات يمكن أن تكون:

Q = أطنان الخرشوف

P = سعر الوحدة من الخرشوف (بالقروش)

Y = متوسط الدخل العائلي السنوي (بآلاف الدولارات)

W = معدل الأمطار السنوية بالبوصات.

تلاحظ الاختلافات بين المعاملات المقطرة في (790) و (791) - وتشير معادلة م ص م إلى أن الطلب على الخرشوف أكثر استجابة للتغيرات في سعره وفي التدخل العائلي عنه مقارنة بتقديرات معادلة م ص ج - وفي ضوء العدد الصغير من المشاهدات في عيشتنا، فلن نتوافر لدينا ثقة كافية، في حال مثل هذه، بأن التقديرات التي حصلنا عليها عن طريق م ص م يمكن الاعتماد عليها. إلا أنه، في حالة النتائج المبنية على عينة كبيرة من القيم المشاهدة للمتغيرات (وبالطبع على بيانات فعلية وليست افتراضية)، فسيكون هناك سبب جيد، بسبب عدم الاتساق في مقدرات م ص ج لتفضيل النتائج المبنية على طريقة م ص م.

نموذج للمالية العامة المحلية

بعد أن رأينا كيف تعمل م ص م في وجود بيانات افتراضية، نتحول الآن لتطبيق هذه الطريقة في دراسة تطبيقية فعلية. والمشكلة التي سندرسها ليست مشكلة مهمة للاقتصاديين، فقط، بل ولرجال الإدارات المحلية، أيضا، وتنحصر هذه المشكلة في التساؤل التالي: ما تأثير السياسات المالية المحلية على قرارات الأفراد؟ هل تؤدي المعدلات العالية لضرائب الملكية إلى تثبيط انتقال الأفراد والمؤسسات المحتمل توطينهما بالمنطقة؟ ما درجة اهتمام المقيمين (الفعلية والمحتملة) بنوعية المدارس المختلفة؟ هل تؤثر نوعية المدارس هذه على قراراتهم بالتوطن في مناطق معينة؟ هذه، بالطبع، ليست أسئلة يسهل الإجابة عنها. ولكنها، ولأسباب واضحة، يهتم بها المسؤولون المحليون في المناطق المختلفة.

ومنذ عشر سنوات، قدم تشارلز تايبوت Charles Tiebout نموذجا نظريا للمالية المحلية يتعامل مع بعض هذه القضايا. * افترض تايبوت نظاما يتكون من عدد كبير من المحليات يقدم كل منها منتجات مختلفة من الخدمات العامة. ويقوم المستهلكون

* يرجع إلى: Charles Tiebout. "A Pure Theory of Local Expenditure.", *Journal of Political Economy*, 64, (Oct. 1956), pp. 416-424.

المتقلون من منطقة لأخرى باختيار المجتمع الذي سيتوطنون به وفقا لتفضيلاتهم لهذه الخدمات. فعلى سبيل المثال فالأفراد ذوو الطلب المرتفع على التعليم سوف يتجمعون في المناطق التي تتوافر بها المدارس المختلفة. وإحدى السمات الجيدة لنموذج تايوت هو أنه يزودنا بالآلية التي يمكن من خلالها أن يعبر الأفراد عن طلباتهم على الخدمات المحلية وإشباعها من خلال قراراتهم بالتوطن (أو كما يطلق عليه أحيانا التصويت بالأقدام "Voting with their feet").

ولكن، على المستوى التطبيقي، يظل التساؤل حول ما إذا كان الأفراد (أو بعضهم، في الأقل) يتصرف بهذه الطريقة أم لا؟. يمكن تخيل الصعوبات العديدة التي تمنع الحركة الكاملة في نموذج تايوت: تكاليف الانتقال من منطقة لأخرى، مواطن العمل وتكاليف المواصلات اليومية. . . وهلم جرا. وعلى الرغم من ذلك، فعند النظر إلى هيكل مناطقنا الحضرية وحركيتها، فإن سلوك الأفراد لا يتعد كثيرا عما يفترضه نموذج تايوت: فالأفراد الذين يعملون في المدن الرئيسية غالبا ماتتاح لهم فرص كبيرة في التوطن عند اختيار الضواحي، وهنا، قد تكون نوعية المدارس المحلية مهمة كثيرا في اختيار مكان الإقامة.

ولكن، كيف نختبر وجود مثل هذا السلوك؟ إحدى طرق ذلك هي التساؤل عن ماذا يتوقع أن نلاحظ إذا كان سلوكا من نوع تايوت له أهمية؟ فإذا كانت نوعية المدارس وعبء الضرائب، في الحقيقة، عوامل مهمة في قرارات الأفراد بالتوطن، فقد نتوقع أن ذلك سينعكس على قيم الملكيات بالمناطق المختلفة. على سبيل المثال، فإن مجتمعا به مدارس ممتازة (مع بقاء الأشياء الأخرى على ما هي عليه) سيكون مرغوبا للتوطن فيه بدرجة أكبر مقارنة بغيره من المجتمعات. وعند محاولة الأفراد الإقامة في ذلك المجتمع فسيدفعون قيم الملكيات إلى أعلى مقارنة بنظيراتها في الأماكن الأخرى. وبالمثل، فإن معدلات عالية من الضرائب، مع ثبات الأشياء الأخرى، ستؤدي إلى تخفيض قيم الملكيات المحلية. وباختصار، في ظل افتراض تايوت، توقع أن تكشف الاختلافات المالية بين المحليات عن نفسها في شكل الاختلافات في قيم الملكيات بكل منها. يوحي ذلك أننا قد نختبر العلاقة

بين المتغيرات المالية وقيم الملكيات عبر المحليات المختلفة. سنفترض أن هذه العلاقة تأخذ الشكل التالي:

$$V_t = b_0 + b_1 X_{1t} + \dots + b_k X_{kt} + b_{(k+1)} Z_{1t} + \dots + b_{(k+l)} Z_{lt} + b_{(k+l+1)} T_t + u_t, \quad (7.92)$$

حيث إن:

V_t مقياس ما لقيم الملكية المحلية في المجتمع رقم t .
 X_{1t}, \dots, X_{kt} المتغيرات غير المالية التي تؤثر في قيم الملكية المحلية.
 Z_{1t}, \dots, Z_{lt} مستويات الإنتاج لعدد t خدمات العامة.
 u_t الخطأ العشوائي.

وبعد ذلك، وباستخدام البيانات من عينة المحليات، يمكننا تقدير هذه العلاقة لنرى ما إذا كانت هناك أي علاقات منتظمة بين خدماتنا العامة والمتغيرات الضريبية وقيم الملكيات المحلية.

سنصف مثل هذه الدراسة*. باستخدام عينة من ٥٣ بلدية محلية في شمال شرق نيوجرسي** (تقع جميعها في إطار منطقة نيويورك-إ-نصرية). تستلزم هذه الدراسة تقدير معادلة تشبه (7.92). يزودنا التعداد العام للسكان والمساكن في الولايات المتحدة الأمريكية بقدر ضخم من المعلومات المتعلقة بسمات السكان والمساكن في تلك البلديات. وعن طريق تكملة هذه البيانات بالمعلومات المالية حول الانفاق المحلي ومعدلات الضرائب من مدينة نيوجرسي، يمكننا، حيثئذ أن نجمع

* يرجع إلى:

W. E. Oates., "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and The Tiebout Hypothesis" *Journal of Political, Economy*, 77(Nov.-Dec.,

1969), pp. 957-971.

** بالولايات المتحدة الأمريكية.

بعض المقاييس المختلفة للمتغيرات في (7.92). واستخدمت الدراسة: *
 $V_t =$ القيمة الوسيطة للسكن المشغول بمالكه في المحلية رقم t باعتباره
 مؤشرا لقيم الملكيات المحلية.

تعتمد قيمة الوحدات السكنية في مجتمع معين على عدد من المتغيرات غير
 المالية ومنها مدى قرب البلد من المدينة الرئيسة بالولاية والسمات المادية لوحدات
 السكن، وعديد من السمات غير الملموسة كالاختبارات «البيئية» وهذه هي X_{it} في
 المعادلة (7.92). ولقياس تأثيرها، استخدمت الدراسة متغيرات مستقلة: المسافة
 بالأميال بين البلدية وبين مدينة نيويورك M_t والوسيط لعدد الحجرات لكل منزل
 يشغله مالكة في البلدية R_t والنسب المئوية للمساكن في البلدية التي بنيت قبل عام
 ١٩٥٠ م N_t ، مقياسا لعمر المساكن في البلدية، واستخدم دخل الأسرة الوسيط Y_t
 في البلدية متغيرا تقريبا للسمات غير الملموسة لها. والافتراض هنا هو أن الأسر
 ذات الدخل المرتفع سوف تنحو إلى التوطن في البلديات الأكثر «جاذبية»، وهكذا
 يمثل دخل الأسرة الوسيط مقياسا للسمات غير الملموسة للتوطن في البلدية.
 وبالنسبة للمتغيرات المالية، اشتملت الدراسة على معدل الضرائب الفعال أو الحقيقي
 (T) على الملكية المحلية (وهو المعدل الاسمي بعد تصحيحه ليأخذ في الحسبان
 الممارسات المختلفة لتقدير الضريبة)، ومقياسا للخدمات المحلية، استخدام الإنفاق
 لكل تلميذ E_t في المدارس العامة المحلية. ** والخطوة الأولى هي عمل انحدار V_t
 على هذه المجموعة من المتغيرات المستقلة باستخدام طريقة المربعات الصغرى
 العادية، ويتتبع عن هذ الانحدار المعادلة المقدرة. ***

* هذه البيانات لسنة ١٩٦٠ م.

** للحصول على مزيد من المعلومات حول هذه المتغيرات والتحويلات الرياضية المستخدمة في معادلة
 الانحدار الفعلية، انظر المقالة ذاتها.

*** المتغير P يعد النسبة المئوية للأسرة في البلدية التي تحصل على دخل يقل عن ٣٠٠٠ دولارا سنويا.
 ويساعد هذا في تصحيح القصور في متغير الدخل Y_t (انظر ص ص ٩٦٢-٩٦٣ من المقالة).

$$\hat{V} = -21 - 3.6 \log T + 3.2 \log E - 1.4 \log M$$

$$(2.4) \quad (4.1) \quad (2.1) \quad (4.8)$$

$$+ 1.7R + 0.05N + 1.5Y + 0.3P, \quad R^2 = 0.93. \quad (7.93)$$

$$(4.1) \quad (3.9) \quad (8.9) \quad (3.6)$$

نجد أن المتغيرات المستقلة كافة تحتوي على معاملات مقدرة تأخذ الإشارات المتوقعة. ومن نسب t ، نجد أنه إذا وضعنا أيًا من فرضيات العدم بانتفاء وجود علاقة (مثلاً، $H_1: b_i \neq 0; H_0: b_i = 0$) فإننا سنرفضه عند مستوى معنوية 5%. وتظهر هذه النتائج (عند اللوحة الأولى) على أنها تدعم نموذج السلوك الفردي حيث تمارس المتغيرات المالية تأثيراً مهماً على قرارات التوطن: فكلما ارتفعت معدلات الضرائب (مع ثبات العوامل الأخرى) انخفضت قيمة الوحدة السكنية، وعلى العكس، كلما ارتفع حجم الإنفاق لكل تلميذ، ارتفعت قيمة الوحدة السكنية. وسوف يتبين لنا أن الأفراد يرغبون، عادة، في دفع مبالغ أكبر للإقامة في مجتمعات ذات معدلات ضريبية منخفضة، وذات مدارس ممتازة.

ولكن، إذا فكرنا قليلاً في نموذجنا للانحدار وفي منهج التقدير المتبع، فسوف تظهر بعض المشاكل الخطيرة. ذلك أنه، بينما نرغب في اعتبار بعض السمات المادية والبيئية للبلدية متغيرات خارجية، فإن متغيرائنا المالية هي، في الحقيقة، متغيرات داخلية. فعلى سبيل المثال، يعتمد معدل الضرائب، عادة، على حجم الموازنة العامة، وعلى حجم الوعاء الضريبي الذي يعتمد بدوره على قيم الملكية.* وبالمثل يتحدد مستوى الإنفاق على المدارس بمعدلات الضرائب (ومرة أخرى بقيم الملكية) وعلى الدخل والسياسات السكانية الأخرى.** وفي الحقيقة،

* تعد الضرائب على الملكيات أهم مصادر الإيرادات العامة على مستوى المحليات في الولايات المتحدة الأمريكية - ملحوظة المترجم.

** تؤثر معدلات الضرائب في الإنفاق العام لأنها تمثل في الحقيقة «سعرًا» للخدمات العامة، على سبيل المثال، فإن السكان المتمين لمجتمع يتسم بوعاء ضريبي تجاري وصناعي متسع يمكنهم أن يشتروا خدمات تعليمية عند سعر نسبي منخفض، لأن جزءاً كبيراً من الإنفاق على الطلاب يأتي من الضرائب المدفوعة بواسطة مؤسسات الأعمال المحلية. نتوقع أن يختار المقيمون في مثل هذه المنطقة ميزانية تعليمية أكبر من تلك التي يمكن أن يختاروها في حالة غياب مثل هذا الوعاء الضريبي الصناعي - التجاري المتسع.

يمكن للفرد أن يجادل عن حق بأن العلاقة السالبة المشاهدة في (7.93) بين معدلات الضرائب وقيمة الوحدات السكنية تعكس الحقيقة بأنه، مع توافر ملكيات ذات قيم عالية، يمكن تحصيل حجم معين من الإيرادات مع وجود معدلات منخفضة للضرائب. أي أن V التي تحدد T وليس العكس. وهذا يجعل من الضروري أن نأخذ في الحسبان النتائج المترتبة للمعادلتين الإضافيتين في نموذجنا واللذين يأخذان الشكل العام التالي:

$$T_t = f(V_t, E_t, \dots), \quad (7.94)$$

$$E_t = g(T_t, Y_t, \dots). \quad (7.95)$$

حيث تشير المعادلة (7.94) إلى أن معدل الضرائب دالة في مستوى الإنفاق المحلي وبين المتغيرات الأخرى (مثل V_t) التي تحدد حجم الوعاء الضريبي بينما تبين المعادلة (7.95) أن الإنفاق على التعليم العام دالة في معدل الضرائب والدخل وبعض المتغيرات الأخرى التي تعكس السمات الملائمة لسكان المجتمع رقم t . وباختصار، فإن متغيراتنا المالية [في هذه الحال، لوغاريتم (T_t) ولوغاريتم (E_t)] متغيرات داخلية في نموذجنا، ويعني هذا أنهما، على الأرجح، مرتبطان بالأخطاء العشوائية، ونتيجة لذلك، فإن معاملاتنا المقدرة في المعادلة (7.93) عرضة لتحيز المعادلات الآتية، ولذا، فإن م ص ع ليست بالطريقة الملائمة لتقدير نموذج الانحدار.

لهذا السبب، تعيد الدراسة تقدير المعادلة باستخدام م ص م. ويعني ذلك أنه ينبغي علينا أن نعمل انحدارا للمتغيرين الداخليين ($E_t = (\log E_t)$ و $T_t' = (\log T_t)$) على المتغيرات المحددة مسبقاً في النموذج. وبهذه الطريقة، يتم إيجاد المتغيرات الخالية من المشكلة ($\hat{E}_t = (\log \hat{E}_t)$ و $\hat{T}_t' = (\log \hat{T}_t)$)، ثم نقدر بعد ذلك المعادلة

* ينبغي أن يلاحظ القارئ أن المتغيرات الخالية من مشكلة التحيز ليست هي $\log(\hat{E}_t)$ و $\log(\hat{T}_t)$. أي أننا لانجري انحدار المرحلة الأولى لـ \hat{E}_t و \hat{T}_t ، ثم نأخذ بعد ذلك اللوغاريثمات بدلاً من ذلك نعامل كلاً من $\log E_t$ و $\log T_t$ باعتبارها متغيرات محولة (أي تأخذ شكلاً رياضياً آخر) E_t' و T_t' على المتغيرات المحددة مسبقاً على \hat{E}_t' و \hat{T}_t' . هذه النقطة المهمة يجب مراعاتها لأننا سنين في الفصل الثامن أن استخدام $\log(\hat{E}_t)$ و $\log(\hat{T}_t)$ سوف يؤدي إلى مقدرات متحيزة.

الأولية باستخدام \hat{E}_t و \hat{T}_t بدلا من $\log E_t$ و $\log T_t$. وهكذا نجد أن الدراسة قد استخدمت إحدى السمات المميزة لـ م ص م. تذكر من مناقشتنا السابقة أن استخدام م ص م لا يتطلب تحديدا كاملا للمعادلات الأخرى في النموذج، أي أنه ليس من الضروري أن نحصل على بيانات عن المتغيرات المستقلة كافة في المعادلة (7.94) والمعادلة (7.95)، بل نحتاج فقط لمعلومات عن بعضها. وبالتحديد، فإنه لكي يعمل منهجنا في التقدير، فإنه ينبغي علينا أن نحصل على متغيرين محددين مسبقا بالإضافة إلى المتغيرات الموجودة في المعادلة (7.93). ومنهج الدراسة هو عزل بعض المتغيرات الإضافية المحددة مسبقا التي ستدخل في المعادلتين (7.94) و (7.95). هذه هي متغيرات تؤثر في معدلات الضرائب وفي الإنفاق على المدارس، ومن أمثلتها الملكيات الصناعية والتجارية لكل نسمة والمستويات التعليمية للسكان ونسبة السكان الملتحقين بالمدارس، وهلم جرا.*

بعد تكوين انحدار $\log E_t$ و $\log T_t$ على هذه المتغيرات المحددة مسبقا، ثم حساب $\widehat{\log E_t}$ و $\widehat{\log T_t}$ ، تقدر الدراسة انحدار المرحلة الثانية للمعادلة والحصول على:

$$\begin{aligned} \hat{V} = & -29 - 3.6 \log T + 4.9 \log E - 1.3 \log M + 1.6R \\ & (2.3) \quad (3.1) \quad (2.1) \quad (4.0) \quad (3.6) \\ & + 0.06N + 1.5Y + 0.3P. \\ & (3.9) \quad (7.7) \quad (3.1) \end{aligned} \quad (7.96)$$

ومن المدهش أن نلاحظ في حالتنا هذه (على العكس من المثال السابق) أن المعاملات المقدرة في معادلة م ص م هي، عموما، قريبة جدا من تقديرات م ص ع. فجميعها لها الإشارات نفسها، وتقريبا القيم/ونفس نسب t نفسها، والاختلاف المهم الوحيد هو أن المعامل المقدر لمتغير الإنفاق على التعليم أكبر في م ص م منه في

* سيكون هناك، في الواقع، سبعة متغيرات محددة مسبقا مستخدمة للحصول على $\log \hat{E}_t$ و $\log \hat{T}_t$ ، وتظهر المجموعة الكاملة لهذه المتغيرات في الصفحة ٩٦٥ من المقالة المذكورة.

معادلة م ص ع [على الرغم من أن م ص ع هنا (3.2) يقع داخل ٩٥٪ فترة ثقة للمقدر المبين في المعادلة (7.98)] وتدل النتائج المرتبطة بهذه الدراسة خاصة أن عدم الاتساق الموجود في المقدرات بسبب وجود الآنية في معادلات النموذج غير خطير. وفي حالات أخرى، تكون مقدرات م ص ع ، م ص ع مختلفة اختلافا كبيرا. وأخيرا (وعلى أية حال) نجد أن النتائج تعطينا دلائل تتفق مع النموذج الذي يوضح اهتمام الأفراد بالمتغيرات المالية المحلية عند اختيارهم للبلدية التي سيتوطنون بها.

ملحق: الأخطاء العشوائية المرتبطة ذاتيا في نموذج المعادلات الآنية*

ناقشنا في الفصلين السادس والسابع عددا من المشاكل التي تظهر عند التقدير عندما تتحقق الافتراضات الأساسية للنموذج. وكان منهجنا، حيثئذ، هو التعامل مع كل واحدة من هذه المشاكل على حدة، مع محاولة تطوير طرق أكثر أو أقل إقناعا في تعديل طرقنا في التقدير للتعامل معها. ولكن، قد تظهر هذه المشاكل معا في بعض الحالات في نموذج الانحدار نفسه، ومن ثم، يصعب التعامل معها. نحاول في هذا الملحق الاهتمام بمثل هذه الحالة حيث يعاني النموذج، في حالتنا هذه مشكلة نظم المعادلات الآنية مضافا إليها مشكلة الارتباط الذاتي بين الأخطاء العشوائية، ونحاول اكتشاف كيف يمكن تقدير مثل هذا النموذج. ونأمل أن يزودنا هذا ببعض الملاحظات الدقيقة حول الطرق التي يعالج بها الاقتصاديون القياسيون مشكلتين أو أكثر من مشاكل التقدير في آن واحد.

افترض أننا مهتمون بالمعادلات الآنية التالية في نظام المعادلات الآنية التالي:

$$Y_{1t} = b_0 + b_1 Y_{1t} + b_2 Y_{2t} + b_3 Y_{3(t-1)} + u_t, \quad (7A.1)$$

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad -1 < \rho < 1, \quad (7A.1)$$

* تعتمد المناقشات التالية، بدرجات متفاوتة، على مقالة :

Ray C. Fair "The Estimation of Simultaneous Equation Models with Lagged Endogenous Variables and First order Correlated Errors." *Econometrica*, 38 (May 1970), pp. 507-516.

وأیضا، على مقالة J. Phillip Cooper في الموضوع نفسه *Econometrica*, 40(March 1972), pp. 305-310

حيث إن X_{1t} متغير خارجي ، Y_{2t} متغير داخلي $Y_{3(t-1)}$ متغير داخلي مبطأ ، u_t هو الخطأ العشوائي . وعلى العكس من امثلتنا السابقة ، نفترض أن الخطأ العشوائي مرتبط ذاتيا ، كما يظهر من المعادلة (7A.2) حيث إن ε_t خطأ عشوائي موزع توزيعا طبيعيا مستقل ، ومن ثم ، لا يرتبط بجميع المتغيرات الخارجية (وجميع القيم المبطة لها) في النموذج . نفترض ، أيضا ، أن ε_t لا يعاني الارتباط الذاتي $[E(\varepsilon_s \varepsilon_t) = 0 \text{ for } s \neq t]$ ، وله قيمة متوسطة تساوي الصفر $[E(\varepsilon_t) = 0]$ وله تباين ثابت $[E(\varepsilon_t^2) = \sigma^2]$.

وإذا كانت ρ صفرا في المعادلة (7A.2) فإننا سنعد $Y_{3(t-1)}$ متغيرا محددا مسبقا ، ونكمل كما بينا من قبل . ولكن ، لأننا نفترض هنا أن $\rho \neq 0$ فإننا نواجه صعابا إضافية . وبسبب الارتباط الذاتي ، نتوقع أن تكون $Y_{3(t-1)}$ مرتبطة بـ u_t . ولنرى ذلك ، لاحظ أنه طالما أن Y_{3t} متغير داخلي فإن $Y_{3(t-1)}$ ستعتمد ، عموما ، على $Y_{1(t-1)}$ والتي تعتمد ، بدورها ، على $u_{(t-1)}$. ينتج عن ذلك أن $Y_{3(t-1)}$ سوف تعتمد على $u_{(t-1)}$ وهكذا ستكون مرتبطة بـ u_t في ضوء المعادلة (7A.2) . هذا يعني أننا لانستطيع أن نعامل $Y_{3(t-1)}$ باعتباره متغيرا محدد مسبقا . وينبغي أن يكون واضحا تعميم هذه النتيجة وهو أنه إذا كانت الأخطاء العشوائية مرتبطة ذاتيا فإن المتغيرات الداخلية المبطة ستكون مرتبطة بالأخطاء العشوائية عموما .

لاحظ أننا قد نستمر في اعتبار المتغيرات الخارجية متغيرات محددة مسبقا ، لأنها تظل غير مرتبطة بالأخطاء العشوائية . ويمكننا أن نرى ذلك عن طريق تكرار الإبطاء والإحلال لـ u_t في المعادلة (7A.2) والتي تعطينا :

$$u_t = \varepsilon_t + \rho \varepsilon_{t-1} + \rho^2 \varepsilon_{t-2} + \dots \quad (7A.3)$$

ولما كانت u_t تعتمد في النهاية على القيم المبطة والقيم الحالية لـ ε_t ، وطالما أن هذه القيم لـ ε_t هي مستقلة عن المتغيرات الخارجية افتراضا ، فإنه ينتج عن ذلك أن u_t ينبغي أن يكون غير مرتبط بالمتغيرات الخارجية . وهكذا يمكننا في المعادلة (7A.1) أن نفترض أن X_{1t} تكون غير مرتبطة بـ u_t .

يتبعي علينا الآن أن نعدل طريقتنا في التقدير م ص م لتأخذ في الحسبان الارتباط الذاتي. وقد ظهر من مناقشتنا للقصود السابقة أنه يتبعي علينا أن نحصل أولاً على مقدر p ، ثم نحول المعادلة (7A.1) للتخلص من الخطأ العشوائي المرتبط ذاتياً، ثم نكمل م ص م.

ولسوء الحظ، فإن اشتقاق مقدر لـ p ليس عملية بسيطة كما كان عليه الحال في الفصل الخامس، عندما اعتبرنا الارتباط الذاتي بمعزل عن المشاكل الأخرى. وإذا كانت المعادلة (7A.1) تعاني، فقط، الارتباط الذاتي، فإن النهج، كما وضع في الفصل الخامس هو الحصول على مقدرات متسقة $\hat{b}_0, \dots, \hat{b}_3$ بواسطة م ص ع واستخدامها لتقدير π_1 بواسطة $\hat{\pi}_1 = \hat{b}_0 + \hat{b}_1 X_{1t} + \hat{b}_2 Y_{2t} + \hat{b}_3 Y_{3(t-1)}$ ومن \hat{u}_t ومن $\hat{u}_t = Y_{1t} - (\hat{b}_0 + \hat{b}_1 X_{1t} + \hat{b}_2 Y_{2t} + \hat{b}_3 Y_{3(t-1)})$ سوف نحسب \hat{p} من المعادلة (6.45). ولكن، طالما أن المقدرات الناتجة لـ b_0, \dots, b_3 ستكون غير متسقة، ومن ثم، تؤدي، بدورها إلى إيجاد مقدر غير متسق لـ p ، فيتبعي علينا أن نقدر (7A.1) بواسطة م ص م.

ولتنفيذ طريقة م ص م، نلاحظ أننا، بسبب ارتباط $Y_{3(t-1)}$ بـ u_t ، نعامله كما لو أنه متغير داخلي، $Z_t = Y_{3(t-1)}$ ، مثلاً. ولتبسيط الترميز، دع X_t تشير إلى المتغيرات الخارجية كافة (مشملة على X_{1t}) في النظام في الفترة t . حيثذ تتمثل المرحلة الأولى لـ م ص م في الحصول على القيم المصححة أو «المنقاه» لكل من Z_t و Y_{2t} أي \hat{Z}_t و \hat{Y}_{2t} .

والآن يمكننا أن نحصل نموذجياً على \hat{Y}_{2t} عن طريق انحدار Y_{2t} على جميع المتغيرات الخارجية في النظام X_t . * وبالمثل، سنجعل $\hat{Z}_t = \hat{Y}_{3(t-1)}$ عن طريق انحدار Z_t على جميع القيم المبطة للمتغيرات الخارجية X_{t-1} . ولكن لتضمن تحقق الشروط الفنية لـ م ص م نكون \hat{Y}_{2t} و \hat{Z}_t عن طريق عمل انحدار لـ Y_{2t} و Z_t على كل من

* تذكر أنه، بسبب وجود الارتباط الذاتي، يتبعي أن نعامل المتغيرات الداخلية المبطة باعتبارها متغيرات داخلية حالية، لذلك، تكون المتغيرات الخارجية هي المتغيرات المحددة مسبقاً، فقط، في النموذج.

المتغيرات المشار إليها لـ X_t و X_{t-1} (على سبيل المثال، على كل من $X_{1(t-1)}$ و X_{1t} من بين متغيرات أخرى). وعندما نحصل على \hat{Y}_{2t} و \hat{Z}_t فإن اكمال منهجنا ينبغي أن يكون واضحاً. أي أننا سوف نقدر معلمات (7A.1) عن طريق إحلال \hat{Y}_{2t} و $\hat{Y}_{3(t-1)}$ محل Y_{2t} و $Y_{3(t-1)}$ على الترتيب، ونضع نجمة على الخطأ العشوائي ثم نشق المعادلات الطبيعية في المرحلة بطريقتنا العادية.

فإذا كنا مهتمين، فقط، بالحصول على مقدرات متسقة لـ b_0, b_1, b_2, b_3 فقد يمكننا حل معادلاتنا الطبيعية للحصول على $\hat{b}_0, \dots, \hat{b}_3$ ونتوقف عندها. ولكننا نرغب عادة في تحديد التباينات ونسب t المرتبطة بمقدراتنا حتى يمكننا إنشاء فترات الثقة واختبار الفروض. ولسوء الحظ فإننا، إذا توقفنا عند هذه النقطة من منهجنا للتقدير، لا يمكننا الحصول على التباينات أو نسب t ، بسبب استمرار وجود الارتباط الذاتي معنا الذي جعل صيغنا للتباين غير صحيحة. ولذا، نضطر لعمل تعديلات إضافية.

ولما كان استخدامنا لـ m ص م قد أنتج لنا مقدرات متسقة $\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3$ ومعلمات (7A.1)، فيمكننا الحصول على مقدر متسق للخطأ العشوائي u_t من خلال:

$$\hat{u}_t = Y_{1t} - (\hat{b}_0 + \hat{b}_1 X_{1t} + \hat{b}_2 Y_{2t} + \hat{b}_3 Y_{3(t-1)}) \quad (7A.4)$$

وباستخدام منهجنا من الفصل الخامس، نحصل الآن على مقدر متسق لـ ρ :

$$\hat{\rho} = \frac{\sum_{t=2}^n \hat{u}_{t-1} \hat{u}_t}{\sum_{t=2}^n \hat{u}_{t-1}^2} \quad (7A.5)$$

* ينبغي أن يكون القارئ قادراً على إثبات أنه إذا كُوت \hat{Y}_{2t} عن طريق إجراء انحدار Z_t على X_{t-1} ، حيث، عموماً، فإن الشروط الموضوعة في (7.61) لن تتحقق. أي إذا جعلنا $Y_{2t} = \hat{Y}_{2t} + \hat{\theta}_{1t}$ و $Y_{3t} = \hat{Y}_{3t} + \hat{\theta}_{2t}$ حيث، فإن $\Sigma(\hat{\theta}_{1t} \hat{Y}_{2t}) \neq 0$ ، $\Sigma(\hat{\theta}_{2t} \hat{Y}_{3t}) \neq 0$ ، (مساعدة للحل: ستكون $\hat{\theta}_{1t}$ مثل أي متغير خارجي حالي مثل X_{1t} ، $\Sigma(\hat{\theta}_{1t} \hat{X}_{1t}) = 0$ ، ولكن \hat{Z}_t سوف تعتمد على $X_{1(t-1)}$).

وهذا (كما نتذكر) هو مقدر ρ المقترح بوساطة المعادلة (7A.2). وكما اوضحنا في الفصل الخامس نبطئ المعادلة (7A.1) لفترة زمنية واحدة ثم نضرب هذه المعادلة المبטئة بوساطة ρ وبعدئذ نطرح معادلتنا الناتجة من (7A.1) لنحصل على *:

$$Y_{1t}^* = B + b_1 X_{1t}^* + b_2 Y_{2t}^* + b_3 Y_{3(t-1)}^* + \varepsilon_t, \quad (7A.6)$$

حيث إن $B = (b_0 - \hat{\rho}b_0)$ وأن:

$$\begin{aligned} Y_{1t}^* &= Y_{1t} - \hat{\rho}Y_{1(t-1)}, & X_{1t}^* &= X_{1t} - \hat{\rho}X_{1(t-1)}, \\ Y_{2t}^* &= Y_{2t} - \hat{\rho}Y_{2(t-1)}, & X_{3t}^* &= X_{3t} - \hat{\rho}X_{3(t-1)}. \end{aligned}$$

وبعد أن نكون قد أزلنا مشكلة الارتباط الذاتي فعليا، يمكننا الآن أن نكمل التحليل بمثل ما قمنا به سابقا باستثناء وحيد في السابق هو أننا استخدمنا جميع المتغيرات الخارجية (X_t) وقيمها المبטئة (X_{t-1}) لتكوين القيم المصححة لـ Y_{2t} و $Y_{3(t-1)}$ ، ولكننا الآن مهتمين بالمتغيرات الداخلية Y_{2t}^* و $Y_{3(t-1)}^*$ في (7A.6). ** هذه المتغيرات (على الترتيب) هي مركبات خطية من Y_{2t} و $Y_{3(t-1)}$ وقيمها المبטئة لفترة واحدة. لذلك، ففي محاولة أخرى لتحقيق التوافق للمتغيرات، نبني \hat{Y}_{2t}^* و $\hat{Y}_{3(t-1)}^*$ عن طريق إجراء انحدار لـ \hat{Y}_{2t}^* ومثلا $Z_t^* = Y_{3(t-1)}^*$ وعلى توليفاتها الخطية المناطرة من X_t و X_{t-1} على وجه التحديد $X_t^* = (X_t - \hat{\rho}X_{t-1})$ و $X_{t-1}^* = (X_{t-1} - \hat{\rho}X_{t-2})$. وعلى سبيل المثال، فإن أحد المتغيرات الذي نرمز له بـ X_t^* سيكون $X_{1t}^* = (X_{1t} - \hat{\rho}X_{1(t-1)})$. وحينما نحسب كلا من \hat{Y}_{2t}^* و $\hat{Y}_{3(t-1)}^*$ نقوم بإحلالها في (7A.6) محل Y_{2t}^* و $Y_{3(t-1)}^*$ ، ونضع نجمة على الخطأ العشوائي ε_t ، ثم نشق بعد ذلك معادلتنا الطبيعية للمرحلة الثانية من الانحدار. ويمكن إثبات أن (في ظل تحقق شروط عامة) المقدرات الناتجة تكون متسقة ولها تباينات العينة الكبيرة المعطاة بوساطة الصيغ

* مرة أخرى، تكون المعادلة (7A.6) صحيحة تماما بدلالة الاحتمالات، فقط، في حالة ما إذا كانت العينة ذات حجم لانهايتي. وسبب ذلك (بالطبع) هو أن $\hat{\rho}$ هو، فقط، مقدر متسق لـ ρ .

** ينبغي علينا أن نعد باستمرار متغيرا داخليا بسبب وجود المشاكل في استخدام $\hat{\rho}$ بدلا من المعلمة الحقيقية ρ في (7A.6).

المعتادة. لذا، يمكننا أن نكون قرات الثقة واختبار الفرضيات بالطريقة المعتادة. لاحظ أن مقدرنا b_0 سيكون $\hat{b}_0 = \hat{B}/(1-\hat{\rho})$. وبالمثل سيصبح تباين العينة الكبيرة لـ \hat{b}_0 :

$$\text{var}(\hat{b}_0) = \frac{1}{(1-\hat{\rho})^2} \text{var}(\hat{B}).$$

نلخص الآن نتائجنا ونعممها: إذا أعطينا معادلة في نظام من المعادلات الآتية تحتوي على أخطاء عشوائية مرتبطة ذاتيا، فإن المتغيرات الداخلية المستقلة المبطة في تلك المعادلة ينبغي أن تعامل باعتبارها متغيرات داخلية، وتقدر مثل هذه المعادلة عن طريق : أولا، إجراء انحدار المتغيرات الداخلية المستقلة الحالية والمتغيرات المستقلة الداخلية المبطة على جميع المتغيرات الخارجية في النموذج وعلى قيمها المبطة لبناء القيم المقدرة ثم نحل بعد ذلك المتغيرات المقدرة محل المتغيرات المستقلة الحالية والمبطة، ونكمل كالعادة للحصول على مقدرات متسقة للمعاملات في معادلة الانحدار. وباستخدام هذه المقدرات لـ م ص م، نوجد متسق للخطأ العشوائي الذي يستخدم، بدوره، للحصول على مقدر لـ ρ ، $\hat{\rho}$ ، في نظام الارتباط الذاتي، ويمكننا، حينها، أن نحول معادلتنا الأصلية للتخلص من الارتباط الذاتي. ويستلزم الأمر استخدام م ص م مرة ثانية، مع تذكر أننا هنا، أيضا، ينبغي أن نعامل المتغيرات المستقلة الداخلية والمبطة باعتبارها متغيرات داخلية (ليست محددة مسبقا). نكون انحدارا لقيمتنا «المحولة» للمتغيرات الداخلية الحالية والمتغيرات الداخلية المستقلة المبطة على المتغيرات الخارجية الحالية والمتغيرات الخارجية المبطة، وبإحلال المتغيرات المصححة محل المتغيرات المستقلة، نتقدم لتنفيذ المرحلة الثانية لطريقة م ص م.

أسئلة

١- افترض النموذج :

$$C_t = b_0 + b_1 C_{t-1} + b_2 Y_t + \varepsilon_{1t}, \quad (1)$$

$$Y_t = I_t + C_t, \quad (2)$$

$$I_t = a_0 + a_1 Y_t + a_2 Y_{t-1} + a_3 r_t + \varepsilon_{2t}, \quad (3)$$

حيث إن Y, I, C و r هي الإنفاق الاستهلاكي، الاستثماري، الدخل ومعدل الفائدة على الترتيب، افترض أن $\varepsilon_1, \varepsilon_2$ ليسا مرتبطين ذاتيا وأنهما مستقلان عن r_t

(أ) اذكر المتغيرات الداخلية والمتغيرات المحددة مسبقا في النموذج

(ب) كيف تقدر المعادلة (1) ؟

(ج) كيف تقدر المعادلة (3) ؟

٢- افترض النموذج التالي لسلوك الأجور - الأسعار :

$$\dot{W}_t = a_0 + a_1 (UN)_t + a_2 \dot{P}_t + \varepsilon_{2t},$$

$$\dot{P}_t = b_0 + b_1 \dot{M}_t + b_2 (UN)_t + b_3 \dot{W}_t + \varepsilon_{1t},$$

حيث إن :

$$\dot{W} = \text{نسبة التغير في الأجور،}$$

$$UN = \text{معدل البطالة،}$$

$$\dot{P} = \text{نسبة التغير في الأسعار،}$$

$$\dot{M}_t = \text{نسبة التغير في عرض النقود،}$$

$$\varepsilon_2, \varepsilon_1 = \text{الأخطاء العشوائية}$$

افترض أن ε_{1t} و ε_{2t} لها متوسطات صفرية، وتباينات ثابتة وليست مرتبطة ذاتيا ومستقلة عن $(UN)_t$ وعن \dot{M}_t .

(أ) هل المعادلات السابقة مميزة ؟ وضح.

(ب) اذكر الخطوط العامة لطريقة تقدير المعادلة المميزة.

٣- افترض النموذج :

$$L_t = a_0 + a_1 W_t + a_2 S_t + u_{1t}, \quad (1)$$

$$W_t = b_0 + b_1 L_t + b_2 P_t + u_{2t}, \quad (2)$$

حيث :

$$L = \text{كمية العمل المستخدم،}$$

W = معدل الأجر،

S = المبيعات،

P = مقياس لإنتاجية العمل.

(أ) أوجد معادلات الشكل المختزل لكل من L_t و W_t .

(ب) وضح الخطوط العريضة لطريقة تقدير المعادلة (1).

٤- افترض أن طلب أحد الأفراد على الأحذية يأخذ الشكل :

$$D_{it} = a_0 + a_1 P_t + a_2 D_{i(t-1)} + u_{it}, \quad (1)$$

حيث D_{it} هو طلب الفرد رقم i على الأحذية في الفترة t ، P_t هو سعر الأحذية، افترض أن :

$$u_{it} = \rho u_{i(t-1)} + \varepsilon_{it}, \quad -1 < \rho < 1,$$

حيث ε_{it} له متوسط صفر، وتباين ثابت، وليس مرتبطاً ذاتياً، ومستقل عن P_t وعن جميع قيمه المعطاة.

(أ) بين أن المتغير التابع المبطأ $D_{i(t-1)}$ مرتبط بالخطأ العشوائي.

(ب) افترض أن المعادلة (1) ليست جزءاً من نظام المعادلات. بين أنه، على

الرغم من ذلك، يمكن تقديرها باستخدام م ص م.

٥- افترض نموذج الانحدار المتعدد التالي :

$$Y_t = b_0 + b_1 X_{1t} + b_2 X_{2t} + u_{1t}, \quad (1)$$

$$X_{2t} = c_0 + c_1 Y_t + u_{2t}, \quad (2)$$

أثبت أنه، في ظل الافتراضات العادية، $E(X_{2t} u_{1t}) \neq 0$.

٦- افترض نموذج الأجور - الأسعار :

$$\dot{W}_t = a_0 + a_1 \dot{P}_t + a_2 (UN_t) + \varepsilon_{1t}, \quad (1)$$

$$\dot{P}_t = b_0 + b_1 \dot{W}_t + \varepsilon_{2t}, \quad (2)$$

حيث :

\dot{W} = نسبة التغير في الأجور النقدية.

\dot{P} = نسبة التغير في الأسعار.

UN = معدل البطالة

(أ) بين أن طريقة م ص م لن تعمل إذا حاولنا تقدير المعادلة (1).

(ب) هل تفشل طريقة م ص م في العمل، أيضا، إذا حاولنا تقدير المعادلة

(2)؟ وضح.

٧- افترض المعادلة الهيكلية التالية التي تعد جزءا من نظام معادلات آنية :

$$Y_{1t} = b_0 + b_1 X_{1t} + b_2 X_{2t} + b_3 Y_{3t} + u_{1t},$$

حيث إن Y_{1t} , Y_{2t} , و Y_{3t} متغيرات داخلية، X_{1t} متغير محدد مسبقا. افترض أن النظام المتكامل الذي أخذت منه هذه المعادلة يحتوي على عشرة متغيرات إضافية محددة سلفا، ولكن، افترض أنه يتوافر لدينا مشاهدات عن واحد منها، فقط، X_{1t} مثلا.

(أ) هل المعادلة مميزة ؟ لماذا نعم أو لماذا لا ؟

(ب) هل يمكننا تقدير المعادلة بوساطة م ص م ؟ اشرح.

٨- افترض نموذج المعادلتين

$$Y_{1t} = a_1 + b_1 X_t^2 + c_1 Y_{2t} + \varepsilon_{1t}, \quad (1)$$

$$Y_{2t} = a_2 + b_2 X_t + c_2 Y_{1t} + \varepsilon_{2t}, \quad (2)$$

حيث X_t : متغير محدد مسبقا، $\varepsilon_1, \varepsilon_2$ تحقق افتراضاتنا المعتادة.

(أ) هل كلا المعادلتين مميزتان ؟ ولماذا ؟

(ب) اشتق معادلات الشكل المختزل.

(ج) وضح، باختصار، طريقة لتقدير المعادلة الأولى من النموذج السابق.

٩- افترض أن الإنفاق الاستثماري الخاص يأخذ الشكل التالي :

$$I_{it} = a + b_1 r_{it} + b_2 S_{i(t-1)} + u_{it}, \quad i = 1, \dots, N,$$

$$r_{it} = r_t + b_3 I_{it} + \varepsilon_{it},$$

حيث :

I_{it} = الإنفاق الاستثماري للمنشأة رقم i في الفترة الزمنية t ،

r_{it} = معدل الفائدة الذي يجب أن تدفعه تلك المنشأة لتمويل الاستثمارات،

$S_{i(t-1)}$ = مبيعات المنشأة في الفترة $t-1$ ،

r_t = متوسط معدل الفائدة في الاقتصاد القومي للأرصدة القابلة للاستثمار.

نفترض أن هذه المنشآت ذات العدد N كبيرة الحجم، لذلك فإن حجم إنفاقها الاستثماري يؤثر على معدل الفائدة الذي تواجهه. افترض تحقق الشروط المعتادة كافة المرتبطة بكل من u_{it} و ε_{it} . افترض، أيضاً، أنه تتوافق لدينا بيانات مقطعية، فقط:

(أ) ناقش هل تلك المعادلات مميزة أم لا ؟

(ب) أوجد معادلة الشكل المختزل لـ I_{it} ؟

نماذج المعادلات الآتية غير الخطية

كانت نماذج المعادلات الآتية التي نوقشت في الفصل السابع كافة نماذج خطية في المتغيرات الداخلية. ولكن كثيراً من (إن لم يكن معظم) النماذج التي يهتم بها الاقتصاديون في الحياة العملية هي نماذج غير خطية في المتغيرات الداخلية. فـ كثير من النماذج الاقتصادية، على سبيل المثال، تحتوي على الأجور، أي W ، والأسعار، P ، باعتبارها متغيرات داخلية. تشرح هذه النماذج، عادة، الطلب على العمل بدلالة معدل الأجر الحقيقي، W/P . ومن الواضح أن متغير الأجر الحقيقي، W/P ، ينبغي أن يفسر بوساطة النموذج، وأنه يكون متغيراً داخلياً إذا كان كل من W و P متغيراً داخلياً. ومن الواضح، أيضاً، أن W/P ليس دالة خطية في المتغيرات الداخلية.

هناك عديد من الأمثلة الأخرى للمتغيرات الداخلية غير الخطية، فالإيراد الكلي عادة ما يعرف بأنه حاصل ضرب (PQ) حيث P السعر و Q عدد الوحدات المباعة. ومرة أخرى، إذا كان كل من P و Q متغيراً داخلياً. فإن النماذج التي تحتوي على الإيرادات الكلية (ربما باعتبارها عنصراً أساسياً للوصول إلى الأرباح) ينبغي أن تعد نماذج غير خطية. وتوجد اعتبارات مشابهة في حالة النماذج التي تحاول إدخال إجمالي الأجور كحاصل ضرب معدل الأجر، W ، في عدد وحدات العمل التي اشترت في كل فترة زمنية، L . وبالمثل، نلاحظ أن معظم النماذج

الاقتصادية الكلية تفسر بعض المقاييس للمستوى العام للأسعار (على سبيل المثال، مكش الناتج القومي الإجمالي) بالإضافة إلى القيم الحقيقية (المكشمة) والجارية لمعظم المتغيرات الاقتصادية المأخوذة في الاعتبار إن لم يكن كلها. وعلى سبيل المثال، تشرح هذه النماذج، عادة، الناتج القومي الإجمالي بالأسعار الثابتة وبالأسعار الجارية أيضاً. وهكذا، فإن مناقشتنا حتى الآن تبين أن النماذج التي تحتوي على كل من القيم الجارية والقيم الحقيقية (المكشمة) للمتغيرات الداخلية، ينبغي أن تعد نماذج غير خطية في المتغيرات الداخلية، إذا كان المكش السعري داخلياً.

وهناك أمثلة إضافية أخرى للنماذج غير الخطية التي تنشأ من دوال الإنتاج غير الخطية في العوامل الداخلية للإنتاج، ومن منحنيات فليس Phillips التي تصاغ بدلالة مقلوب معدل البطالة المحدد داخلياً، وأخيراً، من طبيعة تعريف كثير من المتغيرات الاقتصادية التي يبحث الاقتصاديون تفسيرها. فمثلاً يعرف معدل البطالة بأنه نسبة عدد العاملين العاطلين إلى عدد أفراد قوة العمل. فإذا كانت النماذج الكلية تبحث في تفسير حجم قوة العمل بالإضافة إلى عدد العمال العاطلين عن العمل، فإنها ينبغي أن تعد نماذج غير خطية.

سنناقش في هذا الفصل تحليل مثل هذه النماذج، وعلى نحو خاص سوف نوسع تحليلنا لمشكلة التمييز وتطبيق طريقة المربعات الصغرى ذات المرحلتين على النماذج التي تجسد الاشكال غير الخطية للمتغيرات الداخلية، ولكنها خطية في المعلمات.* وفي ملحق هذا الفصل، سوف نوسع نتائجنا لتشمل تلك المرتبطة

* المرجع الكلاسيكي لمشكلة التمييز في كلا النماذج القياسية الخطية وغير الخطية هو :

F. Fisher. *The Identification Problem in Econometrics*. New York : McGraw-Hill, 1966.

أما مناقشتنا نحن لمشكلة التمييز فستعتمد بشدة على مقالة :

H. H. Kelejjan, "Identification of Nolinear Systems : An Interpretation of Fisher." *Princeton University, Econometric Research Program*, Research Paper No. 22 (Revised), 1970.

وأخيراً، تعتمد مناقشتنا لطريقة المربعات الصغرى ذات المرحلتين على مقالة :

H. H. Keljian. "Two Stage Least Squareis and Econometric Models Linear in Parameters but Nonlinear in the Endogenous Variables." *Journal of American Statitital Association*, June 1971, vol. 66, pp. 373-374

بتقدير النماذج غير الخطية في المعلمات. ويجب أن نحذر القارئ من أن غير الخطية هذه تدخل تعقيدات إضافية في التحليل. ونتيجة لذلك، فإن بعض أجزاء هذا الفصل قد تستدعي درجة أكبر من التركيز، ولكن يشتق التحليل مباشرة وكلية من المادة العلمية الموجودة في الفصول السابقة. وقد يكون من المدهش رؤية امكانية مواصلة التحليل استناداً إلى هذه المادة العلمية.

(٨-١) الإطار التحليلي

ينبغي علينا أولاً أن نحدد، اصطلاحاً، نوع النموذج غير الخطي الذي سنحلله. وفي سبيل عمل ذلك، فإننا لن نوضح المشكلة الأكثر عمومية ولكن، في المقابل، سنوضح الشكل الأكثر تمثيلاً للواقع. وهذا يعني أن معظم النماذج التي نهتم بها في التطبيقات ينبغي أن تلائم هذا الإطار التحليلي. نبدأ بنموذج مكون من معادلتين ثم نعمم النتائج بعد ذلك.

توضيح : نموذج من معادلتين

اعتبر النموذج التوضيحي التالي :

$$Y_{1t} = a_0 + a_1 Y_{2t} + a_2 [Y_{1t} Y_{2t} / X_{1t}] + a_3 X_{2t} + \varepsilon_{1t}, \quad (8.1)$$

$$Y_{2t} = b_0 + b_1 Y_{1t} + b_2 [(Y_{1t} - 2Y_{2t})^2 e^{X_{3t}}] + b_3 X_{4t} + \varepsilon_{2t}, \quad (8.2)$$

حيث X_{1t}, X_{2t}, X_{3t} و X_{4t} متغيرات خارجية، ε_{1t} و ε_{2t} هي أخطاء عشوائية، و Y_{1t}, Y_{2t} هما المتغيران المفترض أن يفسرهما النموذج (أي أنهما متغيران داخليان). افترض أن ε_{1t} و ε_{2t} مستقلان عن المتغيرات الخارجية X_{1s}, X_{2s}, X_{3s} و X_{4s} لجميع t وأن $E(\varepsilon_{it}) = 0$ و $E(\varepsilon_{it}^2) = \sigma_i^2$ حيث $i=1,2$ افترض، أيضاً، أن ε_{1t} و ε_{2t} مستقلان عن ε_{1s} و ε_{2s} لجميع $t \neq s$ ، لذا فإن الأخطاء العشوائية لاتعاني الارتباط الذاتي. هناك سمتان رئيستان مرتبطتان بالنموذج الموصوف في (8.1) و (8.2) ينبغي ملاحظتهما: الأولى أن النموذج خطي في المعلمات a_0, \dots, a_3 و b_0, \dots, b_3

والثانية أن النموذج غير خطي في المتغيرات الداخلية بسبب المتغيرات التي تظهر بين الاقواس تناظر المعلمتين a_2 و b_2 . وينبغي ملاحظة أن قيمة هذه المتغيرات التي تظهر بين الاقواس يمكن تحديدها بواسطة قيم المتغيرات Y_{1t}, Y_{2t} و X_{1t}, X_{2t} . وبالمقابل، يمكن أن نعد المتغيرات داخل الاقواس دوال معروفة لهذه المتغيرات الأربع. ونعني بالدالة المعروفة تلك الدالة ذات الشكل المعروف التي لا تحتوي على معلمات غير معلومة. على سبيل المثال فإن المتغير الموضوع بين قوسين المناظر لـ b_0 ليس على الشكل $[(Y_{1t} - \gamma Y_{2t}) e^{\gamma_2 X_{2t}}]$ ، حيث إن γ_1 و γ_2 هي معلمتان ذواتا قيم غير معلومة، ولو كان الأمر كذلك فلن يكون النموذج خطيًا في المعلمات.

والآن، افترض أن النموذج - لأي مجموعة معطاة من القيم للمعلمات في (8.1) و (8.2) والتي تتفق مع نظرية ذلك النموذج (مثلا لا يمكن أن يكون الميل الحدي للاستهلاك سالبًا) - يمكن حله للحصول على المتغيرات الداخلية Y_{1t}, Y_{2t} بدلالة المتغيرات الخارجية X_{1t}, \dots, X_{4t} والأخطاء العشوائية $\varepsilon_{1t}, \varepsilon_{2t}$. وقد نستطيع، اعتمادًا على بعض المجموعات من القيم الممكنة للمعلمات، الحصول على حل صريح Explicit بطرق سهلة، على سبيل المثال، بوضع $a_2 = b_2 = 0$. وبالنسبة لبعض المجموعات الأخرى من قيم المعلمات قد نستطيع الحصول، فقط، على حل رقمي. قد نستطيع، مثلاً، اشتقاق القيم العادية لـ Y_{1t} و Y_{2t} التي تناظر مجموعة محددة من القيم العددية للمتغيرات الخارجية والأخطاء العشوائية، فقط. ولكن، في أي من هاتين الحالتين، ستعتمد المتغيرات الداخلية Y_{1t} و Y_{2t} ، كما تظهر بالنموذج، على قيم المتغيرات الخارجية والأخطاء العشوائية. وسنشير إلى هذه الظاهرة الآن بالقول أن حل النموذج للمتغيرات الداخلية يعتمد على المتغيرات الخارجية والأخطاء العشوائية أو أنه دالة فيها.

ولأن حل النموذج (8.1) و (8.2) للحصول على Y_{1t} و Y_{2t} يعتمد جزئيًا على الأخطاء العشوائية فلا يمكننا، عمومًا، أن نفترض أن هذه المتغيرات الداخلية والأرقام

العشوائية مستقلة عن بعضها أو حتى غير مترابطة.* هذه النتيجة واضحة من المناقشة الموجودة في الفصل السابع. خذ الآن المتغير الموجود بين الأقواس والمناظر لـ a_2 نجد أن قيمته تعتمد على Y_{1t} , Y_{2t} , و X_{1t} . ولما كانت قيم Y_{1t} و Y_{2t} تعتمد جزئياً على الخطأ العشوائي فإن قيم هذا المتغير الذي بين الأقواس تعتمد ذاتها على الأخطاء العشوائية. ونتيجة لذلك، فإنها ستكون، عموماً، مرتبطة بالأخطاء العشوائية. وبالمثل، فقد نستنتج ارتباط المتغير الموجود بين الأقواس والمناظر لـ b_2 بالأخطاء العشوائية، وللتعميم، نقرر أن أي دالة لمتغير داخلي واحد أو أكثر سوف ترتبط، عموماً، بالأخطاء العشوائية.

بعض التوضيحات

قبل أن نتقدم في التحليل، ينبغي أن نوضح بعض النقاط المهمة للقراء. افترض أن النموذج يحتوي على عدد M من المتغيرات الداخلية (Y_{m1}, \dots, Y_{mM}) ، وعدد G من المتغيرات الخارجية X_{1t}, \dots, X_{Gt} . افترض، أيضاً، أن النموذج يحتوي على عدد M من الأخطاء العشوائية $\varepsilon_{1t}, \dots, \varepsilon_{Mt}$. افترض (كما في الحالة السابقة) أن النموذج يمكن حله للحصول على المتغيرات الداخلية بدلالة المتغيرات الخارجية والأخطاء العشوائية. ارمز لهذه الحلول كالتالي :

$$\begin{aligned} Y_{1t} &= F_1(X_{1t}, \dots, X_{Gt}, \varepsilon_{1t}, \dots, \varepsilon_{Mt}) \\ &\vdots \\ Y_{Mt} &= F_M(X_{1t}, \dots, X_{Gt}, \varepsilon_{1t}, \dots, \varepsilon_{Mt}). \end{aligned}$$

* بالنسبة للقراء الأكثر دراية بالاقتصاد القياسي، نلاحظ أن المجموعة غير الخطية من المعادلات أكثر من حل واحد. سنكمل مناقشتنا في ظل الافتراض بأنه إذا كان لمعادلات النموذج أكثر من حل واحد، فإننا نستبعدنا جميعاً باستثناء واحد منها وذلك عن طريق بعض القيود (غير المصرح بها) على المتغيرات في النموذج. على سبيل المثال، لا يمكن أن تكون الأسعار سالبة وهلم جرا. وعلى سبيل التوضيح فإن المجموعة المكونة من المعادلتين $x^2 + y^2 = 20$ ، $|x| = 2|y|$ لها أربعة حلول $(y=2, x=4)$ ، $(y=2, x=-4)$ ، $(y=-2, x=4)$ ، و $(y=-2, x=-4)$ ولكن إذا كنا متأكدين من أن كلا من X, Y ينبغي أن تكون موجبة فإن الحل الوحيد للنموذج هو $(y=2, x=4)$.

وعموماً، إذا كان النموذج غير خطي فإن الدوال أعلاه والتي تصف اعتماد المتغيرات الداخلية على الأخطاء العشوائية والمتغيرات الخارجية سوف لن تكون خطية. افترض أن النموذج يحتوي، أيضاً، على $(Y_{1t}^2 + Y_{3t} Y_{5t})$ ، أو بتعمق أكثر على $K(Y_{1t}, \dots, Y_{mt})$ بوصفه متغيراً. حيثُ، فإن قيم حل النموذج لهذه المتغيرات $F_{1t}^2 + F_{3t} F_{5t}$ و $F_i(X_{1t}, \dots, X_{Gt}, \varepsilon_{1t}, \varepsilon_m)$ حيث بسطنا الرموز عن طريق التعبير عن $K(F_{1t}, \dots, F_{mt})$ بالرمز F_{it} ، ومن الواضح أنه في النموذج غير الخطي دوال المتغيرات الداخلية على الأخطاء العشوائية والمتغيرات الخارجية، ولكنها ستكون مرتبطة، عادة، بالأخطاء العشوائية لأن قيمها تتحدد جزئياً بواسطة قيم تلك الأخطاء.

ولأغراض التقدير، فقد نعرف المتغير الداخلي بأنه ذلك المتغير الذي يرتبط بالأخطاء العشوائية. لذلك فسوف نشير إلى المتغيرات المركبة Constructed التي هي عبارة عن دوال في واحد أو أكثر من المتغيرات الداخلية - بأنها دوال في المتغيرات الداخلية - أو، ببساطة، متغيرات داخلية. لاحظ أنه من غير المهم كون دالة المتغيرات الداخلية تحتوي، أيضاً، على متغيرات خارجية أم لا، لأن المتغير المركب سيكون مرتبطاً، عموماً، بالأخطاء العشوائية، فقط، بسبب اعتماده على واحد أو أكثر من المتغيرات الداخلية. وبالمناسبة، نلاحظ أن النموذج المكون من المعادلتين (8.1) و (8.2) تحدد فيه قيم المتغيرات الداخلية الأربعة بدلالة المتغيرات الخارجية والأخطاء العشوائية، ولذلك، فإنه إذا كان يمكن التعبير عن Y_{1t} و Y_{2t} بدلالة المتغيرات الخارجية والأخطاء العشوائية. فإن المتغير المركب $(Y_{1t} Y_2 / X_{1t})$ يمكن التعبير عنه، أيضاً، بدلالة هذه المتغيرات نفسها.

توضيح آخر

في مقابل النموذج (8.1) و (8.2)، افترض النموذج ذا المعادلتين :

$$\log(Y_{1t}) = a_0 + a_1 Y_{2t}^3 + a_2 X_{1t} + \varepsilon_{1t}, \quad (8.3)$$

$$Y_{2t}^3 = b_0 + b_1 [\log(Y_{1t})] + b_2 X_{2t} + \varepsilon_{2t}, \quad (8.4)$$

حيث إن Y_{1t} و Y_{2t} متغيران داخليان، X_{1t} و X_{2t} متغيران خارجيان، ε_{1t} و ε_{2t} الأخطاء العشوائية. افترض، أيضاً، أن الأخطاء العشوائية لها السمات المرغوب فيها كافة المذكورة في النموذج السابق.

قد يظهر هذا النموذج، عند اللمحة الأولى، على أنه نموذج غير خطي، ولكن، لأغراض التقدير فإن هذا النموذج يعد نموذجاً خطياً، ولنرى ذلك، نعرف المتغيرين Z_{1t} و Z_{2t} على النحو التالي :

$$Z_{1t} = \log(Y_{1t}), Z_{2t} = Y_{2t}^3. \quad (8.5)$$

حينئذ، وبدلالة هذه المتغيرات، يمكن التعبير عن النموذج (8.3) و (8.4) على

النحو التالي :

$$Z_{1t} = a_0 + a_1 Z_{2t} + a_2 X_{1t} + \varepsilon_{1t}, \quad (8.6)$$

$$Z_{2t} = b_0 + b_1 Z_{1t} + b_2 X_{2t} + \varepsilon_{2t}. \quad (8.7)$$

في هذا الشكل، يظهر النموذج على شكل نموذج مكون من معادلتين وهو خطي في الملمات وفي المتغيرين Z_{1t} و Z_{2t} . لذا، يمكن تقدير ملماته بالطريقة المستخدمة في الفصل السابق. ونستنتج أنه، لأغراض التقدير، يكون النموذج الموجود في (8.3) و (8.4) نموذجاً خطياً.

يمكن تحويل النموذج الموجود في (8.3) و (8.4) إلى نموذج خطي لأن عدد المعادلات، البالغ اثنين، يعادل عدد المتغيرات الداخلية. تذكر الآن أن النموذج في المعادلتين (8.1) و (8.2) له متغيرات داخلية أربعة. ولأن عدد المتغيرات الداخلية في ذلك النموذج يزيد عن عدد المعادلات، فإنه لا يمكن التعبير عن النموذج بوصفه خطياً في متغيرين داخليين، على سبيل المثال، دعنا نعرف Y_{1t} و Y_{2t} على النحو :

$$Z_{3t} = \left[\frac{Y_{1t} Y_{2t}}{X_{1t}} \right]; \quad Z_{4t} = (Y_{1t} - 2Y_{2t})^2 e^{X_{3t}}. \quad (8.8)$$

عندئذ، إذا كان المطلوب التعبير عن النموذج (8.1) و (8.2) بوصفه نموذجاً من معادلتين بدلالة المتغيرين الداخليين Z_{3t} و Z_{4t} فإن المتغيرين Y_{1t} و Y_{2t} ينبغي أن

يعبر عنهما، أيضاً، بدلالة Z_{3t} و Z_{4t} - ويمكن التعبير عن Y_{1t} و Y_{2t} بدلالة Z_{3t} و Z_{4t} و $(X_{1t}$ و $X_{2t})$ عن طريق حل المعادلات في (8.8) للحصول على Y_{1t} و Y_{2t} . ولكن، إذا تم ذلك فسيصبح أن كلا من Y_{1t} و Y_{2t} غير خطيتين في Z_{3t} و Z_{4t} . لتوضيح ذلك، دعنا نعبر عن هذه العلاقات على النحو التالي :

$$\begin{aligned} Y_{1t} &= g_1(Z_{3t}, Z_{4t}, X_{1t}, X_{2t}), \\ Y_{2t} &= g_2(Z_{3t}, Z_{4t}, X_{1t}, X_{2t}), \end{aligned} \quad (8.9)$$

مع ملاحظة أن الدوال الموجودة (8.9) هي دوال غير خطية في Z_{3t} و Z_{4t} ، حيث، يمكن التعبير عن النموذج الموجود في (8.1) و (8.2) بدلالة Z_{3t} و Z_{4t} على النحو التالي :

$$\begin{aligned} g_1(Z_{3t}, Z_{4t}, X_{1t}, X_{2t}) &= a_0 + a_1 g_2(Z_{3t}, Z_{4t}, X_{1t}, X_{2t}) \\ &+ a_2 Z_{3t} + a_3 X_{2t} + \varepsilon_{1t}, \end{aligned} \quad (8.10)$$

$$\begin{aligned} g_2(Z_{3t}, Z_{4t}, X_{1t}, X_{2t}) &= b_0 + b_1 g_1(Z_{3t}, Z_{4t}, X_{1t}, X_{2t}) \\ &+ b_2 Z_{4t} + b_3 X_{4t} + \varepsilon_{2t}, \end{aligned} \quad (8.11)$$

ومن الواضح أن النموذج الموجود في (8.10) و (8.11) ليس نموذجاً خطياً في المتغيرات الداخلية.

تعميم

لأغراض التقدير، تبين النتائج السابقة أنه إذا كان النموذج يحتوي على عدد من المعادلات مساو لعدد المتغيرات الداخلية فإنه يمكن أن يعد نموذجاً خطياً، بغض النظر عن المظهر الذي قد يوحي به ذلك. أما إذا كان عدد المتغيرات الداخلية أكبر من عدد المعادلات، فإن النموذج لا يمكن، عموماً، اختزاله إلى نموذج خطي. وبدقة أكثر، اعتبر نموذج المعادلات الآتية الذي يحتوي على عدد k من المعادلات، والتي تحدد قيمًا فريدة لكل من متغيراته الداخلية بدلالة المتغيرات الخارجية والأخطاء العشوائية.* افترض أن هذا النموذج نموذج خطي في المعلمات. افترض، أيضاً، أن عدد

* مرة أخرى، فإن النموذج غير الخطي سيحدد تحديًا فريدًا قيم جميع متغيراته الداخلية إذا استبعدت جميع الحلول ماعدا احداها وذلك بسبب القيود المختلفة التي ينبغي أن تحققها تلك المتغيرات.

المتغيرات التي تظهر في النموذج وتعتمد على واحد أو أكثر من المتغيرات اللداخلية k^* . افترض أخيراً أنه لا يمكن التعبير عن أي من المتغيرات k^* بوصفه مركباً خطياً من المتغيرات الأخرى (يعنى أنها لا تعاني تعدد العلاقات الخطية)* حيث أن النموذج غير خطي إذا كانت $k^* > k$. أما إذا كانت $k^* = k$ فإن النموذج يمكن اعتباره نموذجاً خطياً لأغراض التقدير. ولم تهتم بالحالة $k^* < k$ لأن هذه الحالة تناظر نظام المعادلات الزائد التحديد overdetermined system (أي أنه يوجد عدد من المعادلات أكبر من عدد المتغيرات الداخلية). فإذا لم تكن بعض المعادلات في مثل هذه النماذج تكراراً لمعادلات أخرى فإن هذه النماذج لن تكون، عموماً، متسقة داخلياً.**

(٧-٨) مشكلة التمييز

توضيح

اعتبر النموذج التالي ذي المعادلتين

$$Y_{1t} = a_0 + a_1 g(Y_{2t}) + a_2 X_t + \varepsilon_{1t}, \quad (8.12)$$

$$Y_{2t} = b_0 + b_1 Y_{1t} + \varepsilon_{2t}, \quad (8.13)$$

حيث إن $g(Y_{2t})$ دالة غير خطية معلومة في Y_{2t} ، وأن X_t متغير خارجي يفترض أنه مستقل عن الأخطاء العشوائية ε_{1t} و ε_{2t} لجميع قيم t و s . افترض أن الأخطاء العشوائية تحقق جميع الصفات المرغوب فيها. فلها متوسطات صفرية، وتباينات ثابتة، وأخيراً ليست مرتبطة ذاتياً.

* وضع هذا الافتراض من أجل استبعاد التكرارات غير المفيدة، على سبيل المثال، فإنه بدون هذا الافتراض، فإن النموذج الذي يحتوي على Y_{1t} ، Y_{2t} ، $3Y_{2t}$ سيوصف بأنه يحتوي على أربعة متغيرات داخلية. ** على سبيل توضيح بسيط، فإن النظام التالي المكون من معادلتين لمتغير واحد $3+2X=5$ و $X+10=5$ ، ليس متسقاً لأن المعادلة الأولى تتضمن أن $X=1$ بينما المعادلة الثانية تتضمن أن $X=5$.

تبين المناقشة الموجودة في الفصل السابع أن معلمات (8.12) ليست مميزة لأن (8.12) تحتوي على متغير داخلي واحد في الطرف الأيمن من المعادلة ولكن لا تستبعد أي من المتغيرات المحددة مسبقًا ولذلك، وفقًا لما جاء بالفصل السابع، فإن محاولة تقدير (8.12) باستخدام م ص م ستفشل بسبب وجود الارتباط الخطي المتعدد التام في المرحلة الثانية. ولكن النموذج الموجود في (8.12) و (8.13) ليس نموذجًا خطيًا، ولذلك لا تنطبق النتائج المرتبطة بالتمييز هنا، وبالتحديد فسرى أنه في ظل وجود بعض الافتراضات الإضافية العامة تكون (8.12) مميزة.

لرؤية ذلك، افترض أن النموذج المكون من المعادلتين (8.12) و (8.13) يحدد قيم المتغيرات الداخلية (Y_{1t} و Y_{2t}) تحديدًا مفردًا بدلالة المتغير الخارجي (X_t) والأخطاء العشوائية ε_{1t} و ε_{2t} . ولتبسيط الرموز دع :

$$Z_t = g(Y_{2t}). \quad (8.14)$$

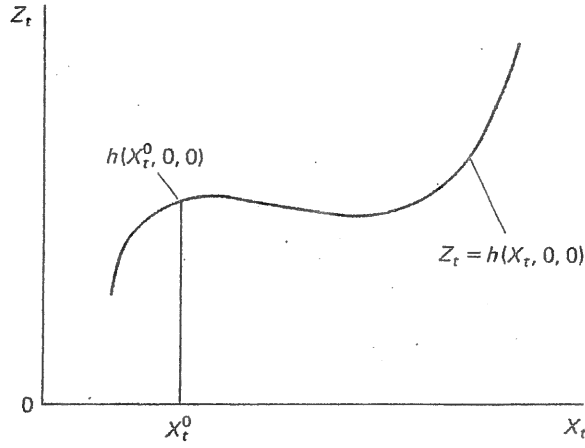
حيث، إذا كان النموذج يحدد قيمة Y_{2t} بدلالة X_t ، ε_{1t} و ε_{2t} فإنه يحدد، أيضًا، قيمة Z_t بدلالة هذه المتغيرات ونرمز لهذا الاعتماد على النحو :

$$Z_t = h(X_t, \varepsilon_{1t}, \varepsilon_{2t}), \quad (8.15)$$

ويلاحظ أنه، طالما أن النموذج غير خطي، فإن الدالة h في (8.15) ستكون، عمومًا، غير خطية. وإذا كانت ε_{1t} و ε_{2t} مساويتين تمامًا (بصفة دائمة) للصفر فإنه يمكن تحديد Z_t كلية بوساطة X_t . لذلك إذا رسمت مشاهدات X_t و Z_t ، فإنه يمكن تتبع منحنى يتوصل إلى معادلته من خلال (8.15) عن طريق وضع $\varepsilon_{1t} = \varepsilon_{2t} = 0$ أي :

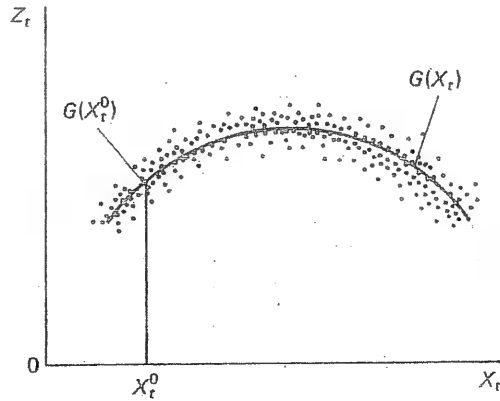
$$Z_t = h(X_t, 0, 0). \quad (8.16)$$

وعلى سبيل التوضيح، فقد تتبعنا المنحنى في الشكل (٨-١) وقد رسم عملاً في شكل غير خطي، لأن النموذج غير الخطي (8.12) و (8.13) يتضمن، عمومًا، أن اعتماد المتغير $Z_t = g(Y_{2t})$ على المتغير الخارجي X_t لن يكون خطيًا.



شكل رقم (٨-١)

نتناول الآن الحالة الأكثر واقعية حيث إن ε_{1t} و ε_{2t} ليستا مساويتين تمامًا ودائمًا للصفر. في هذه الحالة، تتضمن (8.15) أن Z_t لن تحدد كلية بوساطة X_t . ولكن مرة أخرى، بالإشارة إلى (8.15) لن تكون Z_t مستقلة عن X_t لأن X_t هو أحد العناصر المحددة لـ Z_t . ولتوضيح ذلك، افترض أنه توجد لدينا عينة لانهائية من المشاهدات عن X_t و Z_t . حيث إن مناقشتنا تشير إلى أنه إذا رسمنا هذه المشاهدات في شكل بياني فإن شكل الانتشار لهذه النقاط سيكون منحني يعكس الاعتماد الجزئي لـ Z_t على X_t ويظهر مثل هذا الشكل في الشكل رقم (٨-٢).



شكل رقم (٨-٢)

بالنسبة للشكل رقم (٢-٨) نلاحظ أولاً : أن جميع النقاط لاتقع على المنحنى المشار إليه لأن X_t هي ، فقط ، أحد العوامل المحددة لـ Z_t . ثانياً أن هناك عدداً من الطرق التي يمكن للفرد أن يتبع بها منحنى ما بدلالة شكل الانتشار . والمنحنى المشار إليه في شكل رقم (٢-٨) هو المنحنى الذي يعطي القيمة المتوسطة لـ X_t والمناظرة لقيم معطاة لـ X_t . على سبيل المثال ، تكون القيمة المتوسطة لـ Z_t المناظرة لـ $X_t = X_t^0$ هي ارتفاع المنحنى $G(X_t^0)$. ثالثاً : رسمنا عن عمد المنحنى في الشكل رقم (٢-٨) بطريقة مختلفة عن المنحنى الممثل في (١-٨) وسبب ذلك هو أن المنحنى كما حددناه بوساطة شكل رقم (١-٨) سيكون مختلفاً ، عموماً ، عن منحنى العلاقة المتوقعة في شكل رقم (٢-٨) وعلى الرغم من أن ذلك الأمر قد يبدو غير منطقي إلا أنه يمكن توضيحه ، فعلى سبيل المثال ، طبقاً للقيمة المعطاة لـ X_t وهي (X_t^0) تكون القيمة المتوسطة لـ Z_t من (8.15) هي :

$$E(Z_t) = E[h(X_t^0, \varepsilon_{1t}, \varepsilon_{2t})]. \quad (8.17)$$

فإذ كانت الدالة h في (8.17) غير خطية في كل ε_{1t} و ε_{2t} ، فإن نتائجنا في الملحق ب (B) من الفصل الأول وبالتحديد ، (1B.12) تتضمن :

$$E[h(X_t^0, \varepsilon_{1t}, \varepsilon_{2t})] \neq h(X_t^0, E(\varepsilon)_{1t}, E(\varepsilon)_{2t}) = (X_t^0, 0, 0). \quad (8.18)$$

لذلك ، تكون قيمة المنحنى في الشكل (١-٨) التي تناظر $X_t = X_t^0$ [أي $h(X_t^0, 0, 0)$] ليست مساوية للقيمة المناظرة للمنحنى في الشكل (٢-٨) والتي تساوي :

$$G(X_t^0) = E[h(X_t^0, \varepsilon_{1t}, \varepsilon_{2t})]. \quad (8.19)$$

وكما ذكر فإن المنحنى في الشكل رقم (٢-٨) يعطي القيمة Z_t المناظرة لأي قيمة لـ X_t . وبتعريف :

$$V_t = Z_t - G(X_t). \quad (8.20)$$

حيثند ، فإن القيمة المتوسطة V_t المناظرة لأي قيمة معطاة لـ X_t هي الصفر ، على سبيل المثال ، عندما يكون X_t هي X_t^0 ، تكون القيمة المتوسطة لـ V_t كالتالي :

$$E[V_t] = E[Z_t] - G(X_t^0) = G(X_t^0) - G(X_t^0) = 0. \quad (8.21)$$

نتذكر من المبحث (٦-٣) في الفصل السادس ، أنه ، طالما أن القيمة المتوسطة لـ V_t

هي الصفر لأي قيمة معطاة لـ X_t ، فإن القيمة المتوسطة العامة لـ V_t تكون، أيضاً، صفراً، كما أن V_t غير مرتبط بـ X_t .

ويمكن إعادة ترتيب حدود (8.20) على النحو :

$$Z_t = G(X_t) + V_t. \quad (8.22)$$

وفي الحقيقة، فإن العلاقة في (8.22) تشابه، تماماً، تلك العلاقات التي استخدمناها في اشتقاق نماذج الانحدار متعددة الحدود في المبحث (٥-٣). وبالتحديد يمكننا من (8.22) أن نتبين أن القيمة رقم t لـ Z_t مرتبطة ارتباطاً غير خطي بالقيمة رقم t للمتغير المستقل X_t وللحد V_t الذي يمكن اعتباره خطأ عشوائياً، له قيمة متوسطة صفرية تناظر أي قيمة معطاة للمتغير المستقل، وللتوضيح، افترض كما فعلنا في المبحث (٥-٣) أن :

$$G(X_t) \doteq b_0 + b_1 X_t + b_2 X_t^2 + \dots + b_k X_t^k. \quad (8.23)$$

حينئذ، من (8.22)، يكون لدينا :

$$Z_t \doteq b_0 + b_1 X_t + b_2 X_t^2 + \dots + b_k X_t^k + V_t. \quad (8.24)$$

ولتبسيط عرضنا للموضوع، نفترض أن العلاقة بين (8.23) و (8.24) علاقة تساو ونؤكد هنا على أن هذا الافتراض، فقط، من اجل عرض الموضوع، وسيوضح أن النتائج أدناه لا تعتمد على هذا الافتراض.

افترض أن لدينا الآن عدد n من المشاهدات حول متغيرات النموذج (8.12) و (8.13). حينئذ، طالما أن Z_t دالة معلومة في Y_t ، أي $Z_t = g(Y_t)$ فإنه يمكننا أن نحصل على عدد n مشاهدات حول Z_t . ويمكننا، لذلك، اعتبار (8.24) نموذج انحدار، وباستخدام هذا النموذج، تكون القيمة المحسوبة لـ Z_t هي :

$$\hat{Z}_t = \hat{b}_0 + \hat{b}_1 X_t + \hat{b}_2 X_t^2 + \dots + \hat{b}_k X_t^k, \quad (8.25)$$

حيث حصلنا على $\hat{b}_0, \dots, \hat{b}_k$ بوساطة الطريقة المعتادة.

اعتبر الآن مشكلة تقدير (8.12). يكون المتغير الداخلي في الجانب الأيمن هو $g(Y_t)$ ، وهو متغيرنا Z_t . ولتطبيق منهج م ص م للمعادلة (8.12)، يمكننا إحلال القيمة المحسوبة لـ Z_t (أي \hat{Z}_t) محل Z_t . فإذا قمنا بعمل ذلك، فإن المرحلة الثانية

لن تتسم بوجود تعدد العلاقات الخطية، وطالما أن اعتماد \hat{Z}_t على X_t ليس خطياً، فإن \hat{Z}_t لن تكون دالة خطية تامة في X_t . والإيحاء هو أن معلمات (8.12) يمكن أن تقدر باتساق ولذا تكون المعادلات مميزة.

تنقيح

افترضنا، في التحليل السابق، أن دالة القيمة المتوسطة $G(X_t)$ في (8.22) يمكن التعبير عنها في شكل متعدد الحدود في X_t . نين في هذا المبحث أن هذا الافتراض ليس ضرورياً للتمييز، ومن ثم للتقدير المتسق للمعادلة (8.12). وسنن، أيضاً، أنه إذا كان النموذج (8.12) و (8.13) خطياً، بمعنى أن $g(Y_{2t}) = Y_{2t}$ ، فإن المعادلة (8.12) لا يمكن تقديرها باتساق باستخدام طريقة مشابهة للطريقة الموصوفة من قبل وبالتحديد افترض أن $g(Y_{2t}) = Y_{2t}$. افترض، أيضاً، أن Y_{2t} ينحدر على قوى X_t (مثلاً، X_t^k, \dots, X_t^2, X_t) وأن القيمة المحسوبة لـ Y_{2t} يحصل عليها على النحو :

$$\hat{Y}_{2t} = \hat{\alpha}_0 + \hat{\alpha}_1 X_t + \dots + \hat{\alpha}_k X_t^k. \quad (8.26)$$

حيث، فإن تطبيق م ص م لـ (8.12) بعد إحلال \hat{Y}_{2t} محل Y_{2t} يؤدي إلى الحصول على مقدرات غير متسقة.

من الملائم أن نوضح النتيجة الأخيرة في البداية، ثم نوضح بعد ذلك النتيجة الأولى. بداية، نلاحظ أنه إذا كانت $g(Y_{2t}) = Y_{2t}$ ، فإن حل النموذج الخطي (8.12) و (8.13) لـ Y_{2t} (أي معادلة الشكل المختزل) قد يمكن التعبير عنها على النحو التالي :

$$Y_{2t} = \pi_0 + \pi_1 X_t + \psi_t, \quad (8.27)$$

حيث تكون القيمة المتوسطة للمتغير ψ_t لأي قيمة معطاة من X_t هي الصفر $E(\psi_t) = 0$ *. وهكذا (بالنسبة للحالة الخطية)، تكون القيمة المتوسطة لـ Y_{2t} المناظرة

* ينبغي أن يكون واضحاً من المبحث (٧-٤) أن ψ_t ماهي إلا توليفة خطية من الأخطاء العشوائية ε_{1t} و ε_{2t} .

لقيمة معطاة من X_t أي $(\pi_0 + \pi_1 X_t)$. ويتبع عن ذلك أن نموذج الانحدار الذي يربط Y_{2t} بـ X_t ، X_t^2 ، \dots ، X_t^k يمكن التعبير عنه على النحو التالي :

$$Y_{2t} = \pi_0 + \pi_1 X_t + \pi_2 X_t^2 + \dots + \pi_k X_t^k + \psi_t, \quad (8.28)$$

حيث إن $\pi_2 = \pi_3 = \dots = \pi_k = 0$

دع القيمة المحسوبة لـ Y_{2t} من انحدار Y_{2t} على X_t ، X_t^2 ، \dots ، X_t^k هي :

$$\hat{Y}_{2t} = \hat{\pi}_0 + \hat{\pi}_1 X_t + \hat{\pi}_2 X_t^2 + \dots + \hat{\pi}_k X_t^k, \quad (8.29)$$

حيث، من (8.28) ومن افتراضات النموذج (8.12) والنموذج (8.13) يكون لدينا :
 $E(\hat{\pi}_2) = E(\hat{\pi}_3) = \dots = E(\hat{\pi}_k) = 0$ * ويمكن، أيضاً، إثبات أنه، في ظل تحقق بعض الافتراضات الفنية المعقولة فإن $\hat{\pi}_0$ ، $\hat{\pi}_1$ ، \dots ، $\hat{\pi}_k$ هي مقدرات متسقة لمعلماتها المتصلة بها، لذلك، فإن $\hat{\pi}_2$ ، \dots ، $\hat{\pi}_k$ تؤول في الاحتمال إلى الصفر. ويتضمن هذا أنه، إذا كانت $g(Y_{2t}) = Y_{2t}$ وكان حجم العينة لانهائياً فإن \hat{Y}_{2t} سوف تؤول إلى $(\pi_0 + \pi_1 X_t)$ باحتمال قدره الواحد الصحيح. لذلك إذا كان النموذج خطياً، وإذا كان حجم العينة لانهائياً، فإن طريقة م ص م بعد إحلال \hat{Y}_{2t} محل Y_{2t} في (8.29) ستسهم بوجود تعدد العلاقات الخطية التام في المرحلة الثانية وذلك باحتمال قدره الواحد الصحيح. وستفشل الطريقة، ولن يكون المنهج متسقاً. ولكن، لاحظ في حالة العينة النهائية (المحدودة) يمكن تطبيق طريقة م ص م لأن \hat{Y}_{2t} في (8.29) لن يكون مرتبطاً ارتباطاً خطياً تاماً مع X_t . إلا أننا نؤكد هنا أن المنهج ليس متسقاً، لأن خاصية الاتساق ترتبط بوجود عينة لانهائية.

اعتبر الآن الحالة غير الخطية حيث $g(Y_{2t}) \neq Y_{2t}$. وقد أظهرنا، فعلاً، في

هذه الحالة أن $g(Y_{2t})$ يمكن التعبير عنها [انظر (8.22)] على النحو التالي :

$$g(Y_{2t}) = G(X_t) + V_t, \quad (8.30)$$

حيث إن القيمة المتوسطة لـ V_t المناظرة لأي قيمة من قيم X_t تساوي الصفر، وقد أشرنا،

* ينظر النموذج في (8.28) النموذج الخطي العام الذي اعتبرناه في الفصل الرابع. في ذلك الفصل، أوضحنا، بالنسبة لذلك النموذج الذي اعتبرناه، أن مقدرات العلامات تتسم بعدم التحيز.

أيضاً، إلى أن دالة القيمة المتوسطة $G(X_i)$ سوف تكون، عادة، غير خطية في X_i . وإذا كون انحدار لـ $g(Y_{2i})$ على القوى k الأولى من X_i ، وإذا كان حجم العينة لا نهائياً فإن المنحنى المقدر الناتج، أي :

$$\widehat{g(Y_{2i})} = \hat{\alpha}_0 + \hat{\alpha}_1 X_i + \dots + \hat{\alpha}_k X_i^k \quad (8.31)$$

سوف يكون أفضل مقارب لمتعدد الحدود من الدرجة k للمنحنى $G(X_i)$. * أما إذا كانت $G(X_i)$ غير خطية فإن أفضل مقارب لمتعدد الحدود لن يكون خطياً، عموماً. وعادة، يتحسن التقريب بأخذ درجات أعلى لمتعدد الحدود. ويتبع عن ذلك أنه إذا كان حجم العينة لانهائياً فلن يتحول $g(Y_{2i})$ في (8.31)، عموماً، إلى دالة خطية في X_i . ونتيجة لذلك، لا نتوقع أن تفشل طريقة م ص م في حالة العينة اللانهائية.

* للقراء المتخصصين، نرمز إلى متعدد الحدود من الدرجة k في X بالرمز $p(X_i, \alpha)$ حيث إن α ترمز إلى معلماته. حيثذ فإن طريقة المربعات الصغرى سوف توجد α لتدنية :

$$L = \sum_{i=1}^n [g(Y_{2i}) - p(X_i, \alpha)]^2$$

أو، بهدف التوضيح، L/n . ويتذكر أن $V_i = g(Y_{2i}) - G(X_i)$ فإن L/n يمكن التعبير عنها :

$$\begin{aligned} \frac{L}{n} &= \frac{\sum_{i=1}^n [g(Y_{2i}) - G(X_i) + G(X_i) - p(X_i, \alpha)]^2}{n} \\ &= \frac{\sum_{i=1}^n V_i^2}{n} + 2 \frac{\sum_{i=1}^n D_i V_i}{n} + \frac{\sum_{i=1}^n D_i^2}{n} \end{aligned}$$

حيث إن $D_i = G(X_i) - p(X_i, \alpha)$. لاحظ أن قيمة D_i محددة بوساطة قيمة X_i . وبالعودة إلى (8.30) القيمة المتوسطة لـ V_i التي تناظر أي قيمة من قيم X_i هي الصفر. ويتبع عن ذلك أن القيمة المتوسطة لـ V_i والمناظرة لأي قيمة من قيم D_i هي الصفر. ولذلك فتشير مناقشتنا في المبحث (٦-٣) أن V_i غير مرتبطة بـ D_i ، وبسبب ذلك، يمكن إثبات أنه، في ظل تحقيق مجموعة من الافتراضات الإضافية المعقولة، فإن النهاية الاحتمالية لحاصل ضرب التقاطع (cross product) أعلاه يساوي الصفر. لذلك ينبغي أن يكون واضحاً، أنه في حالة العينة اللانهائية فإن L/n تصغر عن طريق اختيار α لتدنيه $\sum D_i^2 / n$ ، طالما أن هذا هو المكون الوحيد لـ L/n الذي يتضمن α .

قاعدة لتمييز النماذج غير الخطية

ينبغي أن يكون القارئ الآن مقتنعاً بأن قواعد التمييز للنماذج الخطية لا يمكن أن تطبق بدون تعديل على النماذج غير الخطية. ونعطي الآن القواعد المناظرة للنماذج غير الخطية، بعدئذ، سنعطي التوضيحات التي تبين أن هذه القواعد معقولة.

افترض نموذجًا مكونًا من عدد M من المعادلات، يكون خطيًا في الملمات لكنه غير خطي في المتغيرات الداخلية، وكل معادلة من هذا النموذج لها علاقة بمتغير اقتصادي معين. ويظهر هذا المتغير، عادة، في الجانب الأيسر من المعادلة، ويكون معاملها ضمنيًا هو الواحد الصحيح.

سنشير إلى مثل هذا المتغيرات بالمتغيرات الداخلية الأساسية. على سبيل المثال تكون المتغيرات الداخلية الأساسية في النموذج (8.12) و (8.13) هي Y_{1t} و Y_{2t} . دعنا نشير إلى جميع المتغيرات الداخلية الأخرى التي تظهر في النموذج بالمتغيرات الداخلية الإضافية. على سبيل المثال، فإن النموذج (8.12) و (8.13) يحتوي على متغير داخلي إضافي واحد وهو $g(Y_{2t})$. تتضمن مناقشتنا في المبحث (٨-١) أنه إذا لم يحتو النموذج على أي متغيرات داخلية إضافية فيعتبر، ولأغراض التقدير، نموذجًا خطيًا.

افترض أن الأخطاء العشوائية للنماذج لها متوسطات صفرية، وغير مرتبطة ذاتيًا، ومستقلة عن جميع قيم المتغيرات الخارجية التي تظهر في النموذج. افترض، أيضاً، أنه يمكن التعبير عن المتغيرات الداخلية الأساسية للنموذج بدلالة الأخطاء العشوائية والمتغيرات الخارجية والمتغيرات الداخلية المبطة (ان وجدت) والمتغيرات الداخلية الإضافية. اعتبر، على سبيل المثال، النموذج الموجود في (8.12) و (8.13). إن شكل المعادلة (8.12) يتخذ الشكل المطلوب حيث إنه لا يظهر في جانبها الأيمن أي من المتغيرات الداخلية الأساسية. ويمكن، بسهولة، الوصول إلى تعبير مشابه لـ Y_{2t} عن طريق التعويض عن قيمة Y_{1t} من (8.12) في (8.15).

أخيراً، افترض أن هناك حلاً وحيداً للنموذج بالنسبة للمتغيرات الداخلية الأساسية بدلالة المتغيرات الخارجية والأخطاء العشوائية والمتغيرات الداخلية المبطة. فإذا لم يتحقق هذا الافتراض فإن النموذج لن يكون كاملاً بمعنى أنه لن يحدد أو يفسر المتغيرات التي بنى النموذج لتفسيرها.

في ظل هذا الافتراضات، وبعض الافتراضات الفنية الأخرى، يمكن إثبات أن معلومات معادلة معينة من النموذج (مثلاً، رقم i) يمكن تقديرها باتساق، ولذلك، تكون تلك المعادلة مميزة إذا كان :

$$A_{1i} \geq A_{2i}, \quad (8.32)$$

حيث إن A_{2i} عدد المتغيرات الداخلية الأساسية التي تظهر في الجانب الأيمن من المعادلة رقم i و A_{1i} عدد المتغيرات المحددة مسبقاً والمتغيرات الداخلية الإضافية التي تظهر في النموذج والتي لا تظهر في المعادلة رقم i ، ويعرب المتغير المحدد مسبقاً في النماذج غير الخطية كالنموذج تحت الدراسة بأنه أي متغير يظهر في النموذج غير مرتبط بالقيم الحالية للأخطاء العشوائية. لذلك، فإن المتغير المحدد مسبقاً سوف يكون أي متغير يظهر في النموذج ولا تعتمد قيمته على القيم المعاصرة لواحد أو أكثر من المتغيرات الداخلية الأساسية (أي أن قيمته ستعتمد، فقط، على المتغيرات الخارجية والمتغيرات الداخلية المبطة).

فإذا أخذنا قاعدة العد في (8.32) بذاتها، فإنها تكون شرطاً ضرورياً لتمييز المعادلة رقم i في النموذج موضوع الاهتمام. ويعني هذا أن، إذا كانت المعادلة رقم i مميزة فإن العلاقة (8.32) ستطبق ولكن العلاقة (8.32) لا تضمن بذاتها أن تكون المعادلة رقم i ، في الحقيقة، مميزة. وفي النماذج غير الخطية، تكون الشروط «الإضافية الفنية» التي ينبغي أن تتحقق لضمان أن المعادلة المعطاة في النموذج مميزة تكون صعبة التحديد، ومن النادر الاهتمام بها في التطبيق ويتضمن المنهج العادي اختبار (8.32) وعند تحقق (8.32)، نفترض أن الشروط «الفنية الإضافية» التي تكون كافية لتحقيق تمييز المعادلة متحققة.

تختلف قاعدة العد المعطاة (8.32) عن تلك التي تقترحها نتائج الفصل

السابع، لأن المتغيرات الداخلية الإضافية تجمع مع المتغيرات المحددة مسبقاً وليس مع المتغيرات الداخلية الأساسية. وقبل محاولة تبرير هذه القاعدة في (8.32)، دعنا نطبق هذه القاعدة بالنسبة للنموذج في (8.12) و (8.13) بالنسبة للمعادلة (8.12) لدينا (بوضع $i=1$)، $A_{11} = 0$ ، طالما لم يتم استبعاد أي من المتغيرات المحددة مسبقاً من المعادلة وأن $A_{21} = 0$ ، وطالما أنه لا تظهر المتغيرات الداخلية الأساسية في الجانب الايمن من المعادلة، لذلك تتحقق (8.32) ($0 \geq 0$) وهكذا بافتراض أن الشروط الفنية الإضافية متحققة تكون (8.12) مميزة، وبالنسبة للمعادلة (8.13) لدينا $A_{12} = 2$ وطالما أن $g(Y_{2t})$ و X_2 مستبعدتان من (8.13)، و $A_{22} = 1$ ، طالما أن Y_{1t} تظهر في الجانب الايمن. لذلك تكون $A_{12} \geq A_{22}$ ، وهكذا مع افتراض تحقق الشروط الفنية الإضافية فإن المعادلة (8.13) مميزة أيضاً.

تبرير القاعدة

دعنا الآن نحاول، أن نوضح لماذا نجمع المتغيرات الداخلية الإضافية في مجموعة واحدة مع المتغيرات المحددة سلفاً لأغراض التمييز. اعتبر مرة أخرى النموذج المبسط في (8.12) و (8.13)، مع ملاحظة أن هذا النموذج نموذج غير خطي بسبب وجود المتغير الداخلي الإضافي $g(Y_{2t})$ ، غير أننا قد أثبتنا، أنه في (8.30)، يمكن التعبير عن $g(Y_{2t})$ باعتباره مجموع مقدارين. الأول هو دالة غير خطية في X_t $[G(X_t)]$ والثاني هو الخطأ العشوائي V_t بقيمة متوسطة تساوي صفراً تناظر أي قيمة لـ X_t وقد بينا أنه يمكن تقريب $G(X_t)$ بدلالة الانحدار متعدد الحدود لـ $g(Y_{2t})$ على قوى X_t [انظر، على سبيل المثال (8.31)].

إذا تم إحلال (8.30) في (8.12) فإن النموذج ذي المعادلتين (8.12) و (8.13) يمكن التعبير عنه على النحو :

$$Y_{1t} = a_0 + a_1 G(X_t) + a_2 X_t + w_t, \quad (8.33)$$

$$Y_{2t} = b_0 + b_1 Y_{1t} + \varepsilon_{2t}, \quad (8.13)$$

حيث إن $w_t = \varepsilon_{1t} + a_1 V_t$. وطالما أن w_t هي مركب خطي من V_t و ε_{1t} فإنه يمكن

اعتباره خطأ عشوائيًا بمتوسط صفر تناظر أي قيمة لـ X_t . لذا، فإن V_t غير مرتبطة بأي من X_t أو $G(X_t)$.

ويمكن اعتبار (8.33) و (8.13) نموذجًا خطيًا مكونًا من معادلتين في المتغيرات التابعة Y_{1t} و Y_{2t} ، والتي تتضمن الحد الثابت $G(X_t)$ و X_t بوصفها متغيرات محددة مسبقًا. لاحظ أن هذا النموذج يحتوي على معلمات الانحدار نفسها كنموذج الأصلي (8.12) و (8.13) (أي a_0, a_1, a_2, b_0, b_1). لاحظ، أيضًا، أنه بخلاف الأخطاء العشوائية، يمكن الحصول على النموذج من النموذج الأصلي عن طريق إحلال المتغير الداخلي الإضافي $G(Y_t)$ بوساطة «دالته المتوسطة» في X_t ، أي $G(X_t)$ ، ومن الواضح أنه إذا كانت (8.33) و (8.13) مميزتين، فإن (8.12) و (8.13) يجب أن تكونا مميزتين أيضًا، حيث إن لهما المعلمات نفسها.

ولأغراض الشرح، سوف تستمر مناقشتنا عن طريق افتراض إمكانية تحديد المشاهدات عن $G(X_t)$ في حالة مشاهدة X_t . هذا الافتراض ليس ضروريًا، إلا أنه يبسط المناقشة. وإحدى الطرق لتبرير هذا الافتراض هي افتراض $G(X_t)$ يمكن تقريبها تقريبًا كاملاً بمتعدد الحدود من الدرجة k في X_t الذي يمكن تقديره باتساق عن طريق انحدار $g(Y_t)$ على قوى X_t . على سبيل المثال، في ظل هذه الافتراضات، ستحدد قيمة $G(X_t)$ من قيمة X_t [انظر (8.31)] على النحو :

$$G(X_t) = \hat{\alpha}_0 + \hat{\alpha}_1 X_t + \dots + \hat{\alpha}_k X_t^k. \quad (8.34)$$

وطالما أننا نعتبر، فقط، حالة العينة الكبيرة ($n = \infty$) فيمكننا أن نسمح «بتقييم كبيرة» جدًا لـ k .

وإذا توافرت المشاهدة عن $G(X_t)$ ، فإن النموذج الموجود في (8.33) و (8.13) يدخل ضمن إطار النماذج الخطية المعتبرة في الفصل السابع. بالتحديد، ومع تحقق شروط فنية إضافية، نرى أن (8.33) مميزة لأن الجانب الأيمن من المعادلة يحتوي، فقط، على متغيرات محددة مسبقًا. وبالمقابل (وبدلالة رموز الفصل السابع)، تكون (8.33) مميزة طالما أن $k_2 \geq k_1$ ، لأن كلا من k_1 و k_2 صفر، حيث إن k_2 هي عدد المتغيرات المحددة مسبقًا والمستبعدة من المعادلة، و k_1 هو عدد المتغيرات

الداخلية في الجانب الأيمن. لاحظ أن هذه النتيجة المرتبطة بـ k_1 و k_2 تكون متماثلة مع النتيجة السابقة المرتبطة بـ A_{11} و A_{21} ويمكن الحصول عليها من المعادلات الأصلية (8.12) و (8.13) من خلال تصنيف المتغير الداخلي الإضافي $g(Y_{2t})$ بوصفه متغيراً محدداً مسبقاً. وبالمثل فيما يتصل بـ (8.13)، يكون لدينا $k_2 \geq k_1$ (طالما أن $k_2 = 2[G(X_{1t})]$ و $k_2 = 2$ مستبعدة) و $k=1$ (طالما أنه Y_{1t} مشتمل عليها) لاحظ، مرة أخرى، أن هذه النتيجة يمكن الحصول عليها عن طريق تصنيف $g(Y_{2t})$ بوصفه متغيراً محدداً مسبقاً في النماذج غير الخطية (8.12) و (8.13).

تعميم لتبرير قاعدة التمييز

يمكن تعميم النتيجة السابقة. ولتوضيح ذلك، نعطي أولاً نتيجة أكثر عمومية تناظر (8.30).

عموماً، قد يحتوي النموذج الخطي المكون من m من المعادلات من الشكل الذي نعالجه على عديد من متغيرات داخلية إضافية، إضافة إلى متغيرات محددة مسبقاً. وفي ظل توفر بعض الافتراضات المعقولة، يمكن التعبير عن كل من هذه المتغيرات الداخلية الإضافية على شكل مجموع لمقدارين إحداهما [يتشابه مع $G(X_{1t})$ السابق] يعطي القيمة المتوسطة للمتغير الإضافي المناظر للقيم المعطاة من المتغيرات المحددة مسبقاً، والآخر هو الخطأ العشوائي ذو القيمة المتوسطة الصفرية لأي مجموعة قيم معطاة للمتغيرات المحددة مسبقاً. وعلى سبيل التوضيح، افترض أن (Y_{1t}, Y_{2t}) متغير داخلي إضافي و X_{1t}, X_{2t}, X_{3t} هي المتغيرات المحددة مسبقاً للنموذج غير الخطي. حيث، وفي ظل افتراضات معقولة، يمكن التعبير عن (Y_{1t}, Y_{2t}) على النحو التالي :

$$(Y_{1t}, Y_{2t}) = H(X_{1t}, X_{2t}, X_{3t}) + \psi_t, \quad (8.35)$$

حيث إن $H(X_{1t}, X_{2t}, X_{3t})$ دالة في X_{1t} و X_{2t} و X_{3t} ، و ψ_t متغير له قيمة متوسطة صفرية مناظرة لأي من قيم المجموعة المعطاة لـ X_{1t} و X_{2t} و X_{3t} . فعلى سبيل المثال، إذا كانت $H(X_{1t}, X_{2t}, X_{3t}) = (X_{1t}^2 + X_{2t})e^{X_{3t}}$ ، وإذا كانت $H_{1t}=3$ ، $X_{2t}=5$ و $X_{3t}=3$

$$H(3,5,0)=(9+5)e^0=14. \quad (8.36)$$
$$H_t = (X_{1t}^2 + X_{2t})e^{x_{3t}} \quad (8.37)$$

إلا أنه خطي في المعلمات، ويجعل المتغيرات الداخلية الرئيسية Y_{11}, \dots, Y_{m1} والمتغيرات الداخلية الإضافية $g_{11}=g_1(Y_{11}, \dots, Y_{m1}), g_{21}=g_2(Y_{11}, \dots, Y_{m1}), \dots, g_{n1}=g_n(Y_{11}, \dots, Y_{m1})$ ، والمتغيرات المحددة مسبقاً X_{11}, \dots, X_{p1} فضمن إطار مناقشتنا أعلاه وفي ظل افتراضات معقولة، يمكن التعبير عن g_{it} على النحو :

$$g_{il} = H_i(X_{1l}, \dots, X_{pl}) + \psi_{il} \quad i = 1, \dots, r, \quad (8.38)$$

حيث إن متوسط ψ_{it} يساوي الصفر لأي مجموعة معطاة من القيم لـ X_{pt}, \dots, X_{it} وبإحلال التعبير المناظر لكل متغير داخلي إضافي محل المتغير الأصلي كما في (8.38) وتجميع الأخطاء العشوائية، بعدئذ، في نهاية الجانب الأيمن من كل معادلة. **

* قد تكون المتغيرات الداخلية الإضافية، عموماً، دوال في المتغيرات المحددة مسبقاً وفي المتغيرات الداخلية الأساسية، أيضاً ومن أجل تبسيط الرموز لم نعالج هذه الحالة.

^{٢٩} على سبيل المثال، افترض أن المعادلة الأولى هي :

$$Y_{1i} = a_0 + a_1 Y_{2i} + a_2 g_{1i} + a_3 g_{2i} + a_4 X_{1i} + \varepsilon_{1i},$$

حيث تتضمن المعادلة (8.38) أنه يمكن إعادة كتابة هذه المعادلة على النحو :

$$Y_{1i} = a_0 + a_1 Y_{2i} + a_2 H_{1i} + a_3 H_{2i} + a_4 X_{1i} + (\varepsilon_{1i} + a_2 \psi_{1i} + a_3 \psi_{2i}),$$

حيث إن $(\varepsilon_1 + a_2\psi_1 + a_3\psi_2)$ يؤخذ على أنه الخطأ العشوائي.

وسوف يحتوي النموذج الخطي الناتج على Y_{1t}, \dots, Y_{mt} بوصفها متغيرات داخلية،
 X_{1t}, \dots, X_{pt} و H_{1t}, \dots, H_{rt} بوصفها متغيرات محددة مسبقًا حيث :

$$H_{it} = H_i(X_{1t}, \dots, X_{pt}), \quad i=1, \dots, r. \quad (8.39)$$

سوف يحتوي هذا النموذج الخطي على معلمات نموذج الانحدار نفسها كما هو الحال بالنسبة للنموذج غير الخطي الأصلي. فإذا طبقت القواعد المرتبطة بالتمييز والمعطاة في الفصل السابع لكل معادلة من هذا النموذج الخطي فستكون النتائج متماثلة مع تلك التي يمكن الحصول عليها عن طريق تطبيق (8.32).

هناك نقطة واحدة ترتبط بالافتراضات «الفنية الإضافية» وقد ذكرنا من قبل أنه قد يكون من المفيد أن نعرض للخطوط العريضة. وفي الحقيقة، فإن النتائج التي توصلنا إليها مبنية على افتراض ضمني هو أن المتغيرات المحددة مسبقًا X_{1t}, \dots, X_{pt} و H_{1t}, \dots, H_{rt} لا تعاني الارتباط الخطي المتعدد، فإذا كانت هذه المتغيرات مرتبطة خطيًا ببعضها بعضًا. فإن نتائجننا في (8.32) لن تكون صحيحة. اعتبر، مرة أخرى، مثلاً، أن النموذج الخطي الذي ينتج عن تطبيق (8.38) للنموذج غير الخطي المشار إليه سابقًا. افترض أن الجانب الأيمن من المعادلة الأولى لهذا النموذج يحتوي على ثلاثة متغيرات داخلية أساسية، مثلاً، Y_{2t} و Y_{3t} و Y_{4t} والحد الثابت والمتغير المحدد مسبقًا X_{1t} . افترض أن هذه المعادلة تستبعد X_{2t} و X_{3t} و H_{1t} ، أيضاً ولكن $H_{1t} = X_{2t} + X_{3t}$. هل يستتج من ذلك أنه إذا تحققت شروط فنية إضافية تكون معادلتنا مميزة؟ والإجابة مطلقاً لا. افترض، مثلاً، أننا نحاول تطبيق طريقة م ص م لتقدير معلمات المعادلة الأولى للنموذج الخطي. سنحاول في المرحلة الأولى أن نحسب \hat{Y}_{2t} ، \hat{Y}_{3t} و \hat{Y}_{4t} عن طريق تكوين انحدار للمتغيرات Y_{2t} ، Y_{3t} و Y_{4t} على الحد الثابت X_{1t} ، X_{2t} ، X_{3t} و H_{1t} . ولكن مجهودات مرحلتنا الأولى سوف تفشل بسبب الارتباط الخطي المتعدد التام بسبب وجود العلاقة الخطية $H_{1t} = X_{2t} + X_{3t}$. ومن الواضح أن هذه المعادلة لن تكون مميزة لأن هناك متغيرين اثنين غير مرتبطين خطيًا قد حذفنا من المعادلة، بينما تحتوي المعادلة على ثلاثة متغيرات داخلية في الجانب الأيمن.

هنا بعض الشروط التي يمكن أن نستنبط على أساسها ما إذا كانت المتغيرات المحددة مسبقاً X_{pt}, \dots, X_{It} و H_{mt}, \dots, H_{It} مرتبطة خطياً مع بعضها بعضاً أم لا. إضافة إلى ذلك فإن النتائج التي عرضناها يمكن تعديّلها لتأخذ في الحسبان هذه الحالة. ولكن المناقشة معقدة. ومثل هذه الحالة من الارتباط الخطي المتعدد التام لاتواجه إلا نادراً في التطبيقات. لذلك ننهي المناقشة بحالة القارئ المتمكن في الاقتصاد القياسي إلى المصادر الأخرى.* ونذكر مرة أخرى قرائنا الآخرين أن تحليلنا ليس شاملاً.

(٨-٣) تقدير م ص م

نعرض في هذا المبحث الخطوط العريضة لطريقة م ص م لتقدير النماذج القياسية الخطية في المعلمات، غير الخطية في المتغيرات الداخلية. والمنهج المقترح هو تعميم مباشر للمنهج الموجود في المبحث (٨-٢).

الخطوط العريضة للطريقة

افترض أن المعادلة رقم i للنموذج القياسي من النوع الذي ندرسه مميزة. حيثئذ، ومع تحقق افتراضات معقولة، يمكن أن نقدر هذه المعادلة على نحو متسق باستخدام الطريقة التالية :

الخطوة الأولى : نحصل على القيم المحسوبة لكل متغير داخلي أساسي يظهر في الجانب الأيمن من المعادلة عن طريق إجراء انحدار لذلك المتغير على

* انظر الفصل الخامس من :

F. Fisher. *The Identification Problem in Econometrics* (New York: McGraw-Hill, 1966), and H. H. Kelejian. "Identification of Nonlinear Systems: An Interpretation of Fisher." *Princeton University, Econometric Research Program, Research Paper No. 22* (Revised), 1970, A nice overall review of the issues involved is given in Chapter 8 of S. Goldfeld and R. Quandt, *Nonlinear Methods in Econometrics* (Amsterdam: North Holland, 1972).

المتغيرات المحددة مسبقاً التي تظهر في النموذج وربما على قوى (أي مربعات أو مكعبات ... الخ) هذه المتغيرات، وسوف نفصل فيما بعد المدى الذي يمكن أن نستخدم فيه قوى المتغيرات المحددة مسبقاً.

الخطوة الثانية : نحصل على القيم المحسوبة للمتغيرات الداخلية الإضافية بالطريقة نفسها الموضحة في الخطوة الأولى.

الخطوة الثالثة : نحل القيم المحسوبة للمتغيرات الداخلية الأساسية والإضافية محل قيمتها في المعادلة رقم i ، وبعدها، نقدر معلمات المعادلة بطريقة المربعات الصغرى.

ففي ظل شروط معقولة، يمكن إثبات أن مقدرات المعلمات التي يحصل عليها في المرحلة الثانية تكون متسقة. والآن نعرض بعض الملاحظات المهمة للمنهج.

ملاحظة (١)

إذا احتوى النموذج على عديد من المتغيرات المحددة مسبقاً، فإن قوى هذه المتغيرات المحددة مسبقاً لن يستخدم في المرحلة الأولى إلا إذا أدى عدم استخدامها في المرحلة الثانية إلى ارتباط خطي تام. والنقطة المهمة هنا هي أن الخاصية المرغوبة فيها لمنهج التقدير (أي خاصية الاتساق) هي خاصية للعينات الكبيرة ($n = \infty$)، فإذا كان حجم العينة لانهائياً فيمكن، في ظل افتراضات معقولة إثبات أن استخدام قوى أعلى للمتغيرات المحددة مسبقاً بالإضافة إلى قواها الدنيا في المرحلة الأولى سيؤدي إلى تباينات أصغر فأصغر للمقدرات التي يحصل عليها في المرحلة الثانية. والمنطق وراء ذلك هو أن انحدارات متعدد الحدود للمرحلة الأولى تصبح تقريبات أفضل وأفضل لدوال القيم المتوقعة المناظرة كلما اخذنا في الاعتبار قوى أعلى. [انظر (8.31) والهامش في صفحة ٣٤٧]. ولكن، طالما أنه عند التطبيق يكون

حجم العينة، عادة، محدوداً فإنه ينبغي أن يكون عدد المتغيرات المستقلة المعتبرة في المرحلة الأولى التي يعتمد جزئياً على عدد القوى المأخوذة في الحساب محدوداً.* وفي الحقيقة فإنه يمكن إثبات أنه إذا اعتبر عدد كبير من القوى للمتغيرات بحيث أصبح عدد المتغيرات في المرحلة الأولى مساوياً لعدد المشاهدات، فإن طريقة المربعات الصغرى ذات المرحلتين تختزل إلى طريقة المربعات الصغرى العادية.** ولذا، يتولد عنها مقدرات غير متسقة. وهنا نقع في معضلة: فلتخفيض تباين العينة الكبيرة، ينبغي زيادة عدد المتغيرات في المرحلة الأولى، ومن الناحية الأخرى، إذا كان حجم العينة محدوداً أي كلما اقترب عدد المتغيرات المستخدمة في المرحلة الأولى عن حجم العينة فإن مقدرات المربعات الصغرى ذات المرحلتين تصبح مشابهة أكثر لمقدرات طريقة المربعات الصغرى العادية غير المتسقة. والنسبة المثلى بين حجم العينة وعدد المتغيرات المستخدمة في المرحلة الأولى موضوع لم يحسم بعد. ولكننا نقترح، إذا كان بالإمكان ذلك أن كون الفرق بين حجم العينة وعدد المتغيرات المستخدمة في المرحلة الأولى ٢٠، في الأقل.

ملاحظة (٢)

اعتبر w عدد المتغيرات المحددة مسبقاً في النموذج، و N هو حجم العينة، حيث، يفترض ضمناً من (١) أن $(N-w) \geq 20$ حيث يقيد في المرحلة الأولى عدد قوى المتغيرات المحددة مسبقاً، فقط، ولكن، في حالة النماذج كبيرة الحجم، قد

* على سبيل المثال، فإن نموذج الانحدار :

$$Y_{1i} = b_0 + b_1 Y_{1i} + b_2 X_{2i} + b_3 X_{2i}^2 + b_4 X_{2i}^2 + \varepsilon_i$$

يحتوي على خمسة متغيرات مستقلة (بما فيها الحد الثابت).

** تتحول م ص م إلى م ص ع إذا كانت القيم المحسوبة للمتغيرات الداخلية غير مساوية القيم الفعلية المناظرة. وفي الحقيقة فإن هذا هو ما يحدث بالضبط، إذا تجاهلنا الدقة النظرية، إذا كان حجم العينة مساوياً عدد المتغيرات بما فيها الحد الثابت في المرحلة الأولى. هذه النتيجة ينبغي أن تكون واضحة. إذا كان هناك قيم وعددها N لمتغير ينبغي أن يفسر بدلالة متغيرات عددها N ، ومن ثم معلومات عددها N ، فإن التوضيح ينبغي أن يكون كاملاً.

تكون من الكبر بحيث تساوي حجم العينة N . أو قد تكون $20 < (N-w)$. في مثل هذه النماذج، ينبغي تقييد عدد المتغيرات المحددة مسبقاً ذات الشكل الخطي التي تدخل في المرحلة الأولى وذلك لأسباب مماثلة لما ذكرنا من قبل في (١). هناك طرق مختلفة عدة لاختيار المتغيرات المحددة مسبقاً التي ستدخل في انحدارات المرحلة الأولى. ولكن، لغرض أن تكون المقدرات متسقة، ينبغي أن تشمل هذه المجموعة من المتغيرات المحددة مسبقاً على جميع المتغيرات المحددة مسبقاً في المعادلة التي نقوم بتقديرها، وينبغي أن تشمل، أيضاً، في الأقل، على عدد مماثل من المتغيرات المحددة مسبقاً غير الموجودة في المعادلة كعدد المتغيرات الداخلية الأساسية والإضافية الموجودة في الجانب الأيمن من تلك المعادلة. وعلى الرغم من أن النموذج الذي ندرسه الآن ليس نموذجاً خطياً، إلا أن مبررات هذه القاعدة هي المبررات نفسها المعطاة في المبحثين (٧-٤) و (٧-٥) من الفصل السابع لحالة النماذج الخطية، ومناقشتنا التالية من الملاحظة الثالثة ستوضح ذلك.

ملاحظة (٣)

ينبغي أن تستخدم مجموعة متغيرات (المرحلة الأولى) المستقلة نفسها في الحصول على المتغيرات المحسوبة كافة التي تستخدم في المرحلة الثانية. وبالتحديد، افترض أن المعادلة رقم i هي :

$$Y_{it} = b_0 + b_1 Y_{1t} + b_2 (Y_{2t} Y_{3t}) + b_3 Y_{2t}^2 + a_1 X_{it} + \varepsilon_{it}. \quad (8.40)$$

افترض، أيضاً، أن النموذج الكامل يحتوي على متغيرات محددة مسبقاً X_{3t} و X_{2t} . دع \hat{Y}_{it} يحصل عليه من الانحدار ^{*}:

$$Y_{it} = \alpha_0 + \alpha_1 X_{1t} + \alpha_2 \cdot '_{2t} + \alpha_3 X_{3t} + \alpha_4 X_{1t}^2 + \alpha_5 X_{2t}^2 + V_{it}. \quad (8.41)$$

دع $Z_{1t} = (Y_{2t} Y_{3t})$ و $Z_{2t} = Y_{2t}$ ، حيثئذ، فإنه لا بد أن يحصل على \hat{Z}_1 من انحدار Z_{1t} على الحد الثابت و X_{1t} ، X_{2t} ، X_{3t} ، و X_{1t}^2 و X_{2t}^2 . وبالمثل، ينبغي أن نحصل

* لاحظ أن (8.41) لا تحتوي على X_{3t}^2 . قد قمنا بذلك، فقط، لتوضيح أنه، بسبب أن انحدار المرحلة الأولى

يحتوي على مربعات X_{1t} و X_{2t} ، فليس هناك حاجة إلى أن يتضمن مربع X_{3t} .

على \hat{Y}_{2t} من انحدار Z_{2t} على المجموعة نفسها من المتغيرات. فإذا لم يستخدم المجموعة نفسها في تحديد \hat{Y}_{1t} ، \hat{Z}_{1t} و \hat{Z}_{2t} فإن المقدرات التي يحصل عليها في المرحلة الثانية لن تكون متسقة. وسوف نشير فيما يلي لسبب ذلك.

ملاحظة (٤)

في وصف طريقة التقدير بالنسبة لـ (8.40)، أشرنا إلى أنه ينبغي في المرحلة الثانية إحلال \hat{Z}_{1t} و \hat{Z}_{2t} محل $Z_{1t} = Y_{2t}Y_{3t}$ و $Z_{2t} = Y_{2t}^2$ على الترتيب. ويمكن إثبات أنه بإحلال $(\hat{Y}_{2t}\hat{Y}_{3t})$ و $(Y_{2t})^2$ محل $(Y_{2t}Y_{3t})$ و Y_{2t}^2 ، حيث يحصل على \hat{Z}_{2t} و \hat{Z}_{3t} بوساطة انحدار Y_{3t} و Y_{2t} على المتغيرات المستقلة للمرحلة الأولى، فإن المقدرات الناتجة لمعاملات الانحدار التي حصل عليها من انحدار المرحلة الثانية لن تكون متسقة. وبشكل أعم، دع $g_{1t} = g_1(Y_{1t}, \dots, Y_{mt})$ متغيراً داخلياً إضافياً يظهر بوصفه متغيراً في المعادلة موضع الاهتمام. حيث، يتطلب التقدير المتسق لمعاملات الانحدار أن يحل \hat{g}_{1t} محل g_{1t} في المرحلة الثانية، ونحصل على \hat{g}_{1t} عن طريق انحدار g_{1t} على المتغيرات المحددة مسبقاً (وربما قوى هذه المتغيرات)، أما إذا تم إحلال g_{1t} في المرحلة الثانية بوساطة $g_1(\hat{Y}_{1t}, \dots, \hat{Y}_{mt})$ حيث سيحصل على كل \hat{Y}_{jt} عن طريق انحدار Y_{jt} على المتغيرات المحددة مسبقاً (وربما قوى هذه المتغيرات) فإن مقدرات معاملات الانحدار لن تكون متسقة. وقبل توضيح المنطق وراء هذا، نعود إلى توضيح لماذا ينبغي استخدام المجموعة نفسها من المتغيرات المستقلة في المرحلة الأولى في تحديد القيم المحسوبة المتغيرات الداخلية كافة.

تبرير لبعض الملاحظات المهمة

اعتبر، مرة أخرى، المعادلة (8.40). افترض أن المتغيرات المحددة مسبقاً للنموذج الذي تكون (8.40) جزءاً منه هي الحد الثابت X_{1t} و X_{2t} . وللتوضيح، افترض أن المتغيرات المستخدمة في المرحلة الأولى هي الحد الثابت X_{1t} ، X_{2t} ، X_{1t}^2 و X_{2t}^2 . ونرمز إلى القيمة المحسوبة لـ Y_{1t} التي يحصل عليها بوساطة انحدار

Y_{1t} على هذه المتغيرات مثل \hat{Y}_{1t} حيث، فإن \hat{Y}_{1t} سوف يكون مولقاً خطياً للمتغيرات المستقلة، مثلاً :

$$\hat{Y}_{1t} = \hat{d}_0 + \hat{d}_1 X_{1t} + \hat{d}_2 X_{2t} + \hat{d}_3 X_{1t}^2 + \hat{d}_4 X_{2t}^2, \quad (8.42)$$

حيث إن $\hat{d}_0, \dots, \hat{d}_4$ هي المعلمات المقدرة في انحدار المرحلة الأولى.

دع $\hat{\phi}_{1t}$ الباقي المقدّر من انحدار المرحلة الأولى، أي :

$$\hat{\phi}_{1t} = Y_{1t} - \hat{Y}_{1t}. \quad (8.43)$$

بعد ذلك، نعلم من نتائجنا في الفصول السابقة أن $\Sigma(\hat{\phi}_{1t}, Y_{1t}) = 0$ طالما أن المعادلات الطبيعية للمرحلة الأولى تكون مبيّنة على الشروط :

$$\begin{aligned} \sum \hat{\phi}_{1t} &= 0, & \sum (\hat{\phi}_{1t} X_{1t}) &= 0, & \sum (\hat{\phi}_{1t} X_{2t}) &= 0, \\ \sum (\hat{\phi}_{1t} X_{1t}^2) &= 0, & \sum (\hat{\phi}_{1t} X_{2t}^2) &= 0. \end{aligned} \quad (8.44)$$

والآن، دع $Z_{1t} = (Y_{2t} Y_{3t})$ ، ودع \hat{Z}_{1t} القيمة المحسوبة لـ Z_{1t} والتي يحصل عليها من انحدار Z_{1t} على المتغيرات المستقلة في المرحلة الأولى نفسها أي الحد الثابت $X_{1t}, X_{2t}, X_{1t}^2, X_{2t}^2$. ودع $\hat{\phi}_{2t}$ الباقي المقدّر المناظر.

$$\hat{\phi}_{2t} = Z_{1t} - \hat{Z}_{1t}. \quad (8.45)$$

نلاحظ أن معادلة الانحدار التي استخدمت لحساب \hat{Y}_{1t} مبيّنة على الشروط :

$$\begin{aligned} \sum \hat{\phi}_{2t} &= 0, & \sum (\hat{\phi}_{2t} X_{1t}) &= 0, & \sum (\hat{\phi}_{2t} X_{2t}) &= 0 \\ \sum (\hat{\phi}_{2t} X_{1t}^2) &= 0, & \sum (\hat{\phi}_{2t} X_{2t}^2) &= 0. \end{aligned} \quad (8.46)$$

نلاحظ، أيضاً، أن الشروط في (8.46) تتضمن أن $\Sigma(\hat{\phi}_{2t}, \hat{Z}_{1t}) = 0$ وأخيراً، دع

$Z_{2t} = Y_{2t}^2$ ودع \hat{Z}_{2t} القيمة المحسوبة لـ Z_{2t} التي حصل عليها عن طريق انحدار Z_{2t} على المجموعة نفسها من المتغيرات المستقلة. دع الباقي المقدّر من هذا الانحدار :

$$\hat{\phi}_{3t} = Z_{2t} - \hat{Z}_{2t}. \quad (8.47)$$

ولاحظ أن $\hat{\phi}_{3t}$ سوف تحقق الشروط :

$$\begin{aligned} \sum \hat{\phi}_{3t} &= 0, & \sum (\hat{\phi}_{3t} X_{1t}) &= 0, & \sum (\hat{\phi}_{3t} X_{2t}) &= 0 \\ \sum (\hat{\phi}_{3t} X_{1t}^2) &= 0, & \sum (\hat{\phi}_{3t} X_{2t}^2) &= 0. \end{aligned} \quad (8.48)$$

لاحظ، أيضاً، أن الشروط في (8.48) تتضمن أن $\Sigma(\hat{\phi}_{3t}, \hat{Y}_{1t}) = 0$ ، إضافة إلى ذلك بسبب أن \hat{Y}_{1t} ، \hat{Z}_{1t} و \hat{Z}_{2t} هي مولفات خطية المجموعة نفسها من المتغيرات، فإن الشروط في (8.44)، (8.46) و (8.48) تتضمن أن :

$$\sum (\hat{\phi}_{it} \hat{Y}_{1t}) = 0, \quad \sum (\hat{\phi}_{it} \hat{Z}_{1t}) = 0, \quad \sum (\hat{\phi}_{it} \hat{Z}_{2t}) = 0 \quad (8.49)$$

حيث $i=1, 2, 3$ ، أي أن مجموع حاصل ضرب التقاطع للبواقي من أحد انحدارات المرحلة الأولى، مع القيم المحسوبة للمتغير من انحدار مرحلة أولى أخرى يساوي الصفر. نحتاج نتيجة أولية إضافية في (8.42)، تربط القيمة المحسوبة لـ Y_{1t} بالمتغيرات المستقلة في المرحلة الأولى عن طريق مقدرات المعلمات $\hat{d}_0, \dots, \hat{d}_4$. فإذا كان حجم العينة لانهائي، فإن هذه المقدرات سوف تؤول إلى ثوابت. دعنا نرمز إلى هذه الثوابت على الترتيب بـ $\hat{d}_0, \dots, \hat{d}_4$ وبالمثل نرمز إلى قيم العينة الكبيرة \hat{Y}_{1t} بالرمز Y_{1t}^m حيث :

$$Y_{1t}^m = d_0 + d_1 X_{1t} + d_2 X_{2t} + d_3 X_{1t}^2 + d_4 X_{2t}^2. \quad (8.50)$$

وبالمثل، دع قيم العينة الكبيرة Z_{1t} و Z_{2t} هي Z_{1t}^m و Z_{2t}^m . دعنا الآن نبين أن اتساق مقدرات م ص م يتطلب مجموعة المتغيرات المستقلة نفسها في انحدارات المرحلة الأولى كافة، يمكن إعادة ترتيب المعادلات (8.43)، (8.45) و (8.47) على النحو :

$$\begin{aligned} Y_{1t} &= Y_{1t}^m + \phi_{1t}, \\ Z_{1t} &= \hat{Z}_{1t} + \hat{\phi}_{2t}, \\ Z_{2t} &= \hat{Z}_{2t} + \hat{\phi}_{3t}. \end{aligned} \quad (8.51)$$

وبالتعويض من المعادلات (8.51) في المعادلة التي نقدرها، أي (8.40)، نحصل على :

$$Y_{it} = b_0 + b_1 \hat{Y}_{1t} + b_2 \hat{Z}_{1t} + b_3 \hat{Z}_{2t} + a_1 X_{1t} + W_t, \quad (8.52)$$

حيث إن $W_t = b_1 \hat{\phi}_{1t} + b_2 \hat{\phi}_{2t} + b_3 \hat{\phi}_{3t} + \varepsilon_{1t}$ ، وبالمثل مناقشتنا في الفصل السابع، نلاحظ أن المكون الوحيد للخطأ العشوائي W_t ذي الصلة هو ε_{1t} طالما أن :

$$\begin{aligned}
\sum W_t &= \sum \varepsilon_{it}, & \sum (\hat{Z}_{2t} W_t) &= \sum (\hat{Z}_{2t} \varepsilon_{it}), \\
\sum (W_t \hat{Y}_{1t}) &= \sum (\hat{Y}_{1t} \varepsilon_{it}), & \sum (X_{1t} W_t) &= \sum (X_{1t} \varepsilon_{it}) \\
\sum (W_t \hat{Z}_{1t}) &= \sum (\hat{Z}_{1t} \varepsilon_{it}).
\end{aligned} \quad (8.53)$$

لذلك، فإن الإيحاء هو تقدير (8.52) عن طريق منهجنا العادي المعدل لطريقة المربعات الصغرى والمعطى بوساطة الشروط التالية [أنظر (8.52)]:

$$\begin{aligned}
\sum \hat{W}_t &= 0, \quad \text{فإن} \quad E \left[\sum \varepsilon_{it} \right] = 0, \quad \text{وطالما أن} \\
\sum (\hat{W}_t \hat{Y}_{1t}) &= 0, \quad \text{فإن} \quad E \left[\sum (Y_{1t}^m \varepsilon_{it}) \right] = 0, \quad \text{وطالما أن} \\
\sum (\hat{W}_t \hat{Z}_{1t}) &= 0, \quad \text{فإن} \quad E \left[\sum (Z_{1t}^m \varepsilon_{it}) \right] = 0, \quad \text{وحيث إن} \\
\sum (\hat{W}_t \hat{Z}_{2t}) &= 0, \quad \text{فإن} \quad E \left[\sum (Z_{2t}^m \varepsilon_{it}) \right] = 0, \quad \text{وحيث إن} \\
\sum (\hat{W}_t X_{1t}) &= 0, \quad \text{فإن} \quad E \left[\sum (X_{1t} \varepsilon_{it}) \right] = 0. \quad \text{وحيث إن}
\end{aligned} \quad (8.54)$$

يمكن إثبات أنه -تحت شروط معقولة- تكون مقدرات الملمات في (8.52) التي يحصل عليها بهذه الطريقة متسقة.

إن عدم استخدام المجموعة نفسها من المتغيرات المستقلة في حساب \hat{Y}_{1t} ، \hat{Z}_{1t} و \hat{Z}_{2t} يعني أن الشروط الموجودة في (8.49) لن تتحقق. ونتيجة لذلك، فإن الشروط كافة في (8.53) سوف لن تتحقق. لذلك، إذا تم تقدير (8.52) بوساطة طريقة المربعات الصغرى العادية المعدلة فإن المقدرات الناتجة سوف تكون مبنية على المعادلات الطبيعية غير المتسقة مع تحديد النموذج. وعلى سبيل التوضيح، افترض أن مبنية على متغيرات مستقلة ليست مماثلة لتلك المستخدمة في \hat{Y}_{1t} و \hat{Z}_{1t} . حينئذ (وعموماً) لا يوجد سبب منطقي في هذه الحالة ليكون مجموع حاصل ضرب التقاطعات $\hat{\phi}_{1t}$ و \hat{Z}_{2t} أو $\hat{\phi}_{2t}$ و \hat{Z}_{2t} مساوياً للصفر وسوف يكون لدينا

(عموماً) $\Sigma(\hat{\phi}_{3t}\hat{Z}_{2t}) \neq 0$ و $\Sigma(\hat{\phi}_{3t}Z_{2t}) \neq 0$ ولذا، وفي هذه الحالة:

$$\Sigma(W_t\hat{Z}_{2t}) = b_1[\Sigma(\hat{\phi}_{1t}\hat{Z}_{2t})] + b_2[\Sigma(\hat{\phi}_{2t}\hat{Z}_{2t})] + \Sigma(\varepsilon_{it}\hat{Z}_{2t}). \quad (8.55)$$

فالمعادلة الطبيعية التي حصل عليها بوساطة وضع $\Sigma(\hat{W}_t\hat{Z}_{2t}) = 0$ (مثلاً) ليست «مقترحة» بوساطة افتراضات النموذج. * على سبيل المثال، تقترح المعادلة (8.55) تقدير معادلتنا بوضع $\Sigma(\hat{W}_t\hat{Z}_{2t}) = b_1[\Sigma(\hat{\phi}_{1t}\hat{Z}_{2t})] + b_2[\Sigma(\hat{\phi}_{2t}\hat{Z}_{2t})]$ ، فإذا فعلنا ذلك فسوف يكون لدينا طريقة تقدير مختلفة وأكثر تعقيداً.

نعود الآن إلى القضية المطروحة في (4). دع $g_t = g(Y_{1t}, \dots, Y_{mt})$ متغيراً داخلياً إضافياً يظهر في المعادلة التي نرغب في تقديرها. سنرى الآن أنه إذا كان طريقة المرحلتين بعد إحلال $g(\hat{Y}_{1t}, \dots, Y_{mt})$ محل $g(Y_{1t}, \dots, Y_{mt})$ في المرحلة الثانية فستكون مقدرات الملاحظات الناتجة غير متسقة.

اعتبر، مرة أخرى، تقدير (8.40)، ولكن افترض الآن أنه قد تم إحلال \hat{Y}_{2t}^2 محل Y_{2t}^2 في المرحلة الثانية، حيث حصل على \hat{Y}_{2t}^2 عن طريق انحدار Y_{2t} على المتغيرات المستقلة في المرحلة الأولى. في هذه الحال، يمكن التعبير عن Y_{2t} على النحو:

$$Y_{2t} = \hat{Y}_{2t} + \hat{\phi}_{4t}, \quad (8.56)$$

حيث إنه من بين أشياء أخرى :

$$\Sigma(\hat{Y}_{2t}\hat{\phi}_{4t}) = 0. \quad (8.57)$$

* نلاحظ بالنسبة للقراء الأكثر دراية، أنه إذا لم نستخدم المجموعة نفسها من المتغيرات المستقلة في جميع انحدارات المرحلة الأولى، فإن البواقي الناتجة عن هذه الانحدارات لن تكون مستقلة إحصائياً عن جميع المتغيرات المستقلة في المرحلة الثانية. وسيؤدي هذا إلى مقدرات غير متسقة للسبب نفسه الذي تؤدي به طريقة م ص م إلى مقدرات غير متسقة في النظم الخطية إذا لم تستخدم المتغيرات المحددة مسبقاً كافة والمتضمنة في المرحلة الأولى.

وبإيجاد مربع جانبي المعادلة (8.56)، يكون لدينا :

$$\begin{aligned} Y_{2t}^2 &= \hat{Y}_{2t}^2 + (\hat{\phi}_{4t}^2 + 2\hat{\phi}_{4t}\hat{Y}_{2t}) \\ &= \hat{Y}_{2t}^2 + \hat{\psi}_t \end{aligned} \quad (8.58)$$

حيث إن $\hat{\psi}_t$ مساو للحد الموجود بين الأقواس في (8.58). ومن الواضح أن $\hat{\psi}_t$ لن تحقق الشروط المشابهة لتلك المعطاة في (8.44)، (8.46) و (8.48). على سبيل المثال في ضوء (8.57) و (8.58) :

$$\sum \hat{\psi}_t = \sum \hat{\phi}_{4t}^2 \neq 0.$$

ينبغي أن يكون واضحاً، أيضاً، أن الخطأ العشوائي في انحدار المرحلة الثانية لن يحقق شروطاً مثل تلك الشروط الموجودة في (8.53). ونتيجة لذلك، فلن تكون المقدرات الناتجة للمعلمات متسقة.

(٨-٤) تبيانات العينة الكبيرة

لحسن الحظ، فإن صيغ التباين المعطاة في الفصل السابع لمقدرات المربعات الصغرى ذات المرحلتين في النظام الخطي للمعادلات ماتزال تنطبق، أيضاً، على المقدرات في النظام غير الخطي. فعلى سبيل المثال، اعتبر، مرة أخرى، التقدير ذا المرحلتين المطبق على (8.40) عن طريق إحلال \hat{Y}_{1t} ، \hat{Z}_{1t} و \hat{Z}_{2t} محل Y_{1t} ، $Z_{1t} = (Y_{2t} Y_{3t})$ و $Z_{2t} = Y_{2t}^2$ على الترتيب. حيثئذ، وفي ظل شروط مماثلة، يمكن إثبات أن مقدرًا متسقًا لتباين العينة الكبيرة \hat{b}_1 سيكون كالتالي :

$$\widehat{\text{var}}(\hat{b}_1) = \frac{\hat{\sigma}_i^2}{\sum \hat{q}_t^2} \quad (8.59)$$

حيث إن $\hat{\sigma}_i^2$ مقدر متسق لتباين ε_{it} و \hat{q}_t الباقي رقم t من انحدار \hat{Y}_{1t} على

الحد الثابت، \hat{Z}_{1t} ، \hat{Z}_{2t} و X_{1t} . والمقدر المتسق الواضح لتباين ε_{it} هو *:

$$\hat{\sigma}_i^2 = \sum_{t=1}^n \frac{(Y_{it} - \hat{b}_0 - \hat{b}_1 Y_{1t} - \hat{b}_2 Z_{1t} - \hat{b}_3 Z_{2t} - \hat{a}_1 X_{1t})^2}{n-5}, \quad (8.60)$$

حيث : n هي حجم العينة، وبالمثل، كما أوضحنا في الفصل السابع، تختبر الفرضيات أو بناء فترات الثقة عن طريق استخدام التوزيع الطبيعي على سبيل التقريب. على سبيل المثال. نجد أن الاستدلالات المتصلة بـ b_1 (في الحالة السابقة) ستكون مبنية على الافتراض بأن:

$$\frac{(\hat{b}_1 - b_1)}{\sqrt{\widehat{\text{var}}(\hat{b}_1)}} \quad (8.61)$$

متغير طبيعي معياري. مرة أخرى، كما في الحالة الخطية، ستكون النتائج صحيحة تمامًا إذا كان حجم العينة لانهائيًا.

(٥-٨) مثال

نعطي الآن مثالاً يوضح ويعمم بعض النتائج التي حصلنا عليها في هذا الفصل. ولأن هدف المثال هو توضيح كيفية الوصول إلى نتائجنا، فلن نهتم بمدى واقعية النموذج أو دقة العلاقات الاقتصادية المتضمنة.

النموذج

اعتبر النموذج الاقتصادي الكلي ذا المعادلات الثلاث :

$$C_t = a_0 + a_1 Y_t + a_2 Y_t^2 + a_3 Y_t^3 + a_4 \left(\frac{1}{C_{t-1}} \right) + a_5 W_{t-1} + u_{1t}, \quad (8.62a)$$

* نلاحظ نقطة مهمة وهي أنه إذا كان حجم العينة لانهائيًا، فإن $n-5 = n = \infty$.

$$I_t = b_0 + b_1(Y_t Y_{t-1})^{1/2} + b_2 r_t + b_3 T_t + u_{2t}, \quad (8.62b)$$

$$Y_t = C_t + I_t + G_t, \quad t = 1, \dots, n, \quad (8.62c)$$

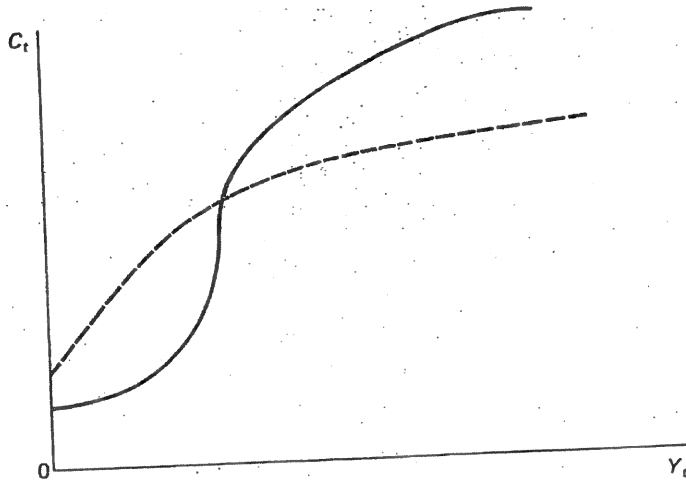
حيث C_t إجمالي الإنفاق الاستهلاكي في الفترة t ، Y_t هو الدخل الإجمالي في الفترة t ، W_{t-1} ثروة المستهلك في الفترة $(t-1)$ ، I_t الإنفاق الاستثماري في الفترة t ، r_t معدل الفائدة في الفترة t ، G_t الإنفاق الحكومي في الفترة t ، T_t متغير اتجاه الزمن Time trend، حيث $T_1 = 1$ و $T_2 = 2$ وهلم جرا، u_{1t} و u_{2t} قيم الأخطاء العشوائية في الفترة t . وقبل أن نحدد افتراضاتنا الإحصائية تحديداً اصطلاحياً سنصف، باختصار طبيعة النموذج.

المعادلة (8.62a) هي دالة الاستهلاك التي تربط الإنفاق الاستهلاكي بالدخل والمستويات السابقة من الاستهلاك (لتعكس أثر العادة) وثروة المستهلك في الفترة الزمنية السابقة. وبالطبع، يتوقع أن يكون للدخل أثر موجب على الاستهلاك، ولكن الشكل الدقيق لهذه العلاقة الموجبة قد لا يكون واضحاً. وبالتحديد، فقد لا تكون هذه العلاقة خطية كما تعرضها، عادة، المراجع الاقتصادية. نعرض في الشكل رقم (٨-٣) علاقيتين محتملتين بين الإنفاق الاستهلاكي والدخل عند قيم مفترضة للمتغيرات الأخرى المتضمنة. يتضمن الشكل المكعب الموجود في (8.62a) مرونة كافية للأخذ في الحسبان كلا من الشكل الخطي ($a_2 = a_3 = 0$) بالإضافة إلى العلاقات الموجودة في الشكل رقم (٨-٣) فإذا أردنا تحقيق مرونة إضافية فيمكننا إضافة حد من الدرجة الرابعة لمتغير الدخل في (8.62).

تفسر المعادلة (8.62b) مستوى الاستثمار بدلالة مستويات الدخل ومعدلات الفائدة ومتغير الاتجاه الزمني. ويمكن أن ننظر إلى المتغير الأخير هذا على أنه صافي مجموع قوى استثمار خارجية يفترض أنها تزيد (إذا كانت $b_3 > 0$) أو تخفض (إذا كانت $b_3 < 0$) بانتظام حجم الاستثمار فترة بعد أخرى. أما الحد $(Y_3 Y_{t-1})^{1/2}$ فهو صياغة معدلة لمبدأ المعجل الذي أدخلناه في التحليل للتوضيح.

وأخيراً تفسر المعادلة (8.62c) الدخل باعتباره مجموع الإنفاقات. وفي

وأخيراً تفسر المعادلة (8.62c) الدخل باعتباره مجموع الإنفاقات. وفي الحقيقة، هذه المعادلة شرط توازني يبين أن الدخل (السلع والخدمات) المنتج المعروض يطلب من المتعاملين في السوق أي من المستهلكين والمستثمرين والحكومة.



شكل رقم (٨-٣)

نفترض أن u_{1t} و u_{2t} غير مرتبطتين ذاتياً، ولهما قيم متوسطة صفرية، (أي $E(u_{1t}) = E(u_{2t}) = 0$) وتباينات ثابتة (أي $E(u_{1t}^2) = \sigma_1^2$, $E(u_{2t}^2) = \sigma_2^2$) وتغاير ثابت $E(u_{1t} u_{2t}) = \sigma_{12}$. نفترض، أيضاً، أن الإنفاق الحكومي وثروة المستهلك ومعدل الفائدة تتولد خارجياً، ولذا، فالأخطاء العشوائية u_{1t} و u_{2t} مستقلة عن G_t و r_t وعن W_{t-1} ولأن متغير الاتجاه الزمني T_t هو متغير يقيني deterministic، كما أنه متغير خارجي، فإن كلا الخطأين العشوائيين مستقلان عنه. وهكذا، فإن نموذجنا في (8.62) نموذج من ثلاث معادلات تفسر كلا من C_t و I_t وأخيراً Y_t بدلالة كل من G_t ، T_t ، r_t ، Y_{t-1} ، W_{t-1} ، C_{t-1} وفي نموذج أعم، قد يحاول الباحث تفسير معدل الفائدة وربما الإنفاق الحكومي من بين أشياء أخرى.

تحليل النموذج

يشتمل النموذج (8.62) على المتغيرات الداخلية الأساسية C_t و I_t وأخيراً Y_t . المتغيرات الداخلية الإضافية هي Y_t^2 و Y_t^3 وأخيراً $(Y_t Y_{t-1})$. يوجد، بالإضافة إلى هذه المتغيرات، حد ثابت وخمسة متغيرات محددة مسبقاً هي $(1/C_{t-1})$ ، W_{t-1} ، r_t ، T_t و G_t . ومن بين هذه المتغيرات الخمسة، نجد أن $(1/C_{t-1})$ هو المتغير الوحيد غير الخارجي حيث يكون النموذج قد حدد في الفترة الزمنية $t-1$ قيمة C_{t-1} ومن ثم، معكوسه، $(1/C_{t-1})$.

اعتبر الآن المعادلة (8.62a). هذه المعادلة تستوفي الشروط الضرورية للتمييز كما توضحها المعادلة (8.32) لاحتوائها على متغير داخلي أساسي واحد، فقط، في الجانب الأيمن وهو Y_t لكنها تستبعد أربعة متغيرات هي إما متغيرات داخلية إضافية أي $(Y_t Y_{t-1})^{1/2}$ أو متغيرات محددة مسبقاً (في هذه الحالة متغيرات خارجية) هي r_t ، T_t و G_t ، كما أن الشرط الضروري لتمييز المعادلة قد تم استوفى لأن هذه المعادلة لا تحتوي على أي متغير داخلي في جانبها الأيمن كما تستبعد خمسة متغيرات تكون إما متغيرات داخلية إضافية هي Y_t^2 و Y_t^3 أو متغيرات محددة مسبقاً هي $1/C_{t-1}$ ، W_{t-1} و G_t ، كما أن قضية التمييز الخاصة بـ (8.62c) لا تنشأ أصلاً لأنها لا تحتوي على معلومات يجب تقديرها.

اعتبر الآن مشكلة تقدير (8.62a) : لتطبيق طريقة م ص م ينبغي أن نحصل أولاً على القيم المحسوبة لـ Y_t ، $Q_{1t} = Y_t^2$ و $Q_{2t} = Y_t^3$. وللحصول على هذه القيم المحسوبة، ينبغي علينا أن نحدد المتغيرات المستقلة لمرحلة الانحدار الأولى، فإذا كانت عيتنا، مثلاً، بحجم $n=50$ فقد تختار المتغيرات المستقلة للمرحلة الأولى

كالتالي :

$$(1/C_{t-1}), W_{t-1}, r_t, T_t, G_t, Y_{t-1}, \quad (8.63)$$

في (8.63)، اخترنا، ببساطة، جميع المتغيرات المحددة مسبقاً في النموذج وقيمها المربعة. لم نضمن قيمها المكعبة أو القوى الأكبر منها أو حاصل ضرب الحدود التقاطعية بينها لأننا نتوقع ارتباط هذه المتغيرات ارتباطاً كبيراً مع المتغيرات الموجودة

في (8.63)، فإذا أدخلنا هذه الحدود الإضافية في (8.63)، فلربما نتوقع أن أي تحسن في المقدرات، بسبب أن الانحدارات متعددة الحدود تقارب بدقة الدوال المتوسطة، سينخفض حجم الخسارة الناتجة عن قلة الفرق بين حجم العينة وعدد المتغيرات المستخدمة في المرحلة الأولى [انظر المبحث (٨-٣)]. ولكن، لا يوجد أي أساس علمي لذلك الاعتقاد! وقد تؤدي إضافة الحدود المكعبة وذات الدرجات الأعلى وحدود حواصل ضرب التقاطعات إلى تحسن في مقدرات معالم الانحدارات.* وعلى أي حال، فإن الاختيار الموجود في (8.63) يشتمل، فعلاً، على المتغيرات المحددة مسبقاً كافة في (8.62a) أي الحد الثابت، $1/C_{t-1}$ ، W_{t-1} ، وفي الأقل، على عدد من المتغيرات المحددة مسبقاً يماثل تلك الموجودة في (8.62a)، أي 10، كما يماثل عدد المتغيرات الداخلية الأساسية والإضافية في (8.62) أي 3. إضافة إلى ذلك، فإن الاختبار في (8.63) يستوفي شرطنا المقترح بأن يكون الفرق بين حجم العينة الذي سيكون 49 بسبب فقد الملاحظة الأولى بسبب المتغيرات المبطأة، وعدد المتغيرات المستخدمة في المرحلة الأولى (أي 13) في الأقل، 20. والآن تتضح باقي الطريقة، حيث يمكننا حساب كل من \hat{Y}_t ، \hat{Q}_{1t} وأخيراً \hat{Q}_{2t} بدلالة انحدارات المربعات الصغرى لـ \hat{Y}_t ، \hat{Q}_{1t} و \hat{Q}_{2t} على المتغيرات في (8.63) حيث $t=2, \dots, n=50$. لاحظ أننا نستخدم المجموعة نفسها من المتغيرات المستقلة الأصلية لحساب \hat{Y}_t ، \hat{Q}_{1t} و \hat{Q}_{2t} . حيث، يكون انحدار المرحلة الثانية المناظر لتقدير (8.62a) كالتالي :

$$C_t = a_0 + a_1 \hat{Y}_t + a_2 \hat{Q}_{1t} + a_3 \hat{Q}_{2t} + a_4 (1/C_{t-1}) + a_5 W_{t-1} + k_t, \quad (8.64)$$

$$t = 2, \dots, n = 50,$$

حيث إن k_t هو الخطأ العشوائي الناتج. وعندئذ، يحصل على مقدرات المعلمات

* إن خاصية الاتساق في مقدرات م ص م هي خاصية للعينة الكبيرة، بينما تكون هذه المقدرات في العينات المحددة متحيزة. لذلك، تعتمد جودة تقدير م ص م على تحيزه وتباينه، وعادة ما يتم جمع هاتين الخاصيتين للمقدر للحصول على ما يسمى متوسط مربع الخطأ mean square error. ولذلك، يمكننا القول، في العينات المحدودة، إن مقدر م ص م لمعلمة معينة «أفضل من» مقدر آخر (والذا قد يكون له مجموعة مختلفة من المتغيرات المستقلة في المرحلة الأولى) إذا كانت له قيمة متوسطة أقل لمربع الخطأ.

، a_0, a_1, a_2, a_3, a_4 و a_5 بدلالة انحدار المربعات الصغرى المناظرة (8.64) ، وبالتحديد، فإن المعادلات الطبيعية لهذا الانحدار معطاة من خلال الشروط التالية:

$$\begin{aligned} \sum_{t=2}^{50} \hat{k}_t &= 0, & \sum_{t=2}^{50} (\hat{k}_t \hat{Y}_t) &= 0, & \sum_{t=2}^{50} (\hat{k}_t \hat{Q}_{1t}) &= 0, \\ \sum_{t=2}^{50} (\hat{k}_t \hat{Q}_{2t}) &= 0, & \sum_{t=2}^{50} [\hat{k}_t (1/C_{t-1})] &= 0, & \sum_{t=2}^{50} (\hat{k}_t W_{t-1}) &= 0, \end{aligned} \quad (8.65)$$

حيث إن :

$$\hat{k}_t = C_t - \hat{a}_0 - \hat{a}_1 \hat{Y}_t - \hat{a}_2 \hat{Q}_{1t} - \hat{a}_3 \hat{Q}_{2t} - \hat{a}_4 (1/C_{t-1}) - \hat{a}_5 W_{t-1}.$$

كما ترمز \hat{a}_i ، حيث $i=0, \dots, 5$ إلى مقدر a_i .

وسيتم تقدير تباين u_{1t} ، σ_1^2 ، بعد ذلك على النحو :

$$\hat{\sigma}_1^2 = \sum_{t=2}^{50} \frac{\hat{u}_{1t}^2}{(49-6)}, \quad (8.66)$$

حيث إن :

$$\hat{u}_{1t} = C_t - \hat{a}_0 - \hat{a}_1 Y_t - \hat{a}_2 Y_t^2 - \hat{a}_3 Y_t^3 - \hat{a}_4 (1/C_{t-1}) - \hat{a}_5 W_{t-1}.$$

وأخيراً، يمكن تقدير تباين العينة الكبيرة لـ \hat{a}_2 (مثلاً) على النحو :

$$\widehat{\text{var}(\hat{a}_2)} = \hat{\sigma}_1^2 \left(\frac{1}{\sum_{t=2}^{50} \hat{Q}_{1t}^2} \right), \quad (8.67)$$

حيث إن \hat{Q}_{1t} هو الباقي من انحدار المربعات الصغرى على الحد الثابت، \hat{Y}_t و \hat{Q}_{2t} ، $(1/C_{t-1})$ وأخيراً W_{t-1} .

وتتمثل خطوات تقدير (8.62b) مع تلك المستخدمة في (8.62a) . والنقطة الوحيدة التي ينبغي ملاحظتها هي أن المتغيرات المستقلة في انحدار المرحلة الأولى في (8.62b) ليست بالضرورة متطابقة مع تلك المستخدمة في تقدير (8.62a) . أي أنها

ليست بالضرورة مطابقة مع المجموعة في (8.63). وبالمقابل ، ويغض النظر عن طريقة اختيار المتغيرات المستقلة في المرحلة الأولى لتقدير معادلة معينة، فإن هذه المجموعة نفسها ينبغي أن تستخدم في تحديد القيم المحسوبة لجميع المتغيرات الداخلية الأساسية والإضافية الموجودة في الجانب الأيمن لتلك المعادلة، وليس من الضروري أن نستخدم المجموعة نفسها لجميع المعادلات. ومن الناحية الأخرى قد يكون المدافع ضعيفاً لتغيير المتغيرات المستقلة في المرحلة الأولى نظراً لتماثل حجم العينة في جميع المعادلات وتماثل المتغيرات.

ملحق: تقدير النماذج غير الخطية في كل من

المتغيرات الداخلية والمعلومات

نوسع تحليلنا في هذا الملحق ليشمل تقدير نماذج المعادلات الآتية غير الخطية في كل من المتغيرات الداخلية والمعلومات. لانهم هنا بقضية التمييز، لأنه لم توجد حتى الآن قواعد بسيطة نسبيًا تستخدم عند التطبيق. وطريقة التقدير التي نعرضها هي واحدة من الطرق التي يمكن فهم مبرراتها النظرية. ومالم يكن الشخص ملماً بالتحليل العددي والبرمجة فإن التطبيق يتطلب توافر برامج حاسوب على مثل هذه الطرق على سبيل الاختيار.*

إطار التحليل

اعتبر نموذجًا مكونًا من ثلاث معادلات يحتوي على المتغيرات الداخلية: Y_{1t} , Y_{2t} و Y_{3t} والمتغيرات الخارجية X_{1t} , ..., X_{kt} . افترض أن المعادلة الأولى من هذا النموذج هي:

$$Y_{1t} = a_0 + a_1 X_{1t} e^{a_2 Y_{2t}} + a_3 X_{2t} + a_4 Y_{3t}^2 + \varepsilon_{1t}, \quad (8A.1)$$

حيث إن ε_t : الخطأ العشوائي. افترض أن قيمة a_2 غير معلومة، ولذا، ينبغي تقديرها بالإضافة إلى قيم a_0 , a_1 , a_3 و a_4 . حيث، وعلى العكس من النماذج كافة التي اعتبرناها حتى الآن، نرى أن معادلة (8A.1) غير خطية في كل من المتغيرات الداخلية وفي المعلومات. لاحظ أنه لا يمكننا أن نحول (8A.1) إلى نموذج خطي في المعلومات عن طريق التعبير عنه على النحو:

$$Y_{1t} = a_0 + a_1 Z_t + a_3 X_{2t} + a_4 Y_{3t}^2 + \varepsilon_{1t}, \quad (8A.2)$$

* هي طريقة المربعات الصغرى غير الخطية ذات المرحلتين، وقد تم اقتراحه أولاً بواسطة T. Amemiya. "The Nonlinear Two-stage Least - Squares Estimator". *Journal of Econometrics* 2(1974), pp. 105-110.

وستكون مناقشتنا هي تفسير نتائج Amemiya.

حيث إن Z_t هو المتغير الداخلي الإضافي ($Z_t = X_{1t} e^{a_2} Y_{2t}$). وسبب ذلك هو أن قيمة a_2 غير معلومة، ولذلك لا يمكننا بناء المشاهدات عن Z_t من تلك المتاحة عن X_{1t} و Y_{2t} . ومثال آخر لنموذج غير خطي في كل من المتغيرات الداخلية وفي الملاحظات هو النموذج المقترح التالي المكون من المعادلتين :

$$\log(Y_{1t}) = a_0 + a_1 \left(\frac{Y_{2t}}{1 + b_2 X_{2t}} \right) + a_2 \left(\frac{X_{1t}}{Y_{2t}} \right) + \varepsilon_{1t}, \quad (8A.3a)$$

$$(Y_{2t}^\alpha X_{3t}) = b_0 + b_1 Y_{1t} + b_2 Y_{1t}^2 + b_3 X_{2t} + \varepsilon_{2t}, \quad (8A.3b)$$

حيث إن Y_{1t} و Y_{2t} هي المتغيرات الداخلية، X_{1t} و X_{2t} و X_{3t} هي المتغيرات الخارجية، و ε_{1t} و ε_{2t} الأخطاء العشوائية. في هذه الحال، تحدث عدم الخطية في الملاحظات بسبب الملاحظات α و b_2 .

نستخدم الآن النماذج في كل من (8A.1) و (8A.3) لتوضيح خاصيتين للنماذج من النوع الذي نعتبره في هذا الملحق. الأولى هي أن الجانب الأيسر من معادلة معينة قد يحتوي على ملاحظات غير معلومة (كما في (8A.3b)) ولكن هذا ليس ضروريًا كما في (8A.1) و (8A.3a). الثاني يكون الخطأ العشوائي في كل معادلة من النوع الذي يضاف إلى بعضه بعضًا. وسوف نشير بعد ذلك إلى أن هذا الافتراض ليس مقيدًا جدًا. فإذا قبلنا هذا الفرض مؤقتًا نلاحظ أن ذلك يتضمن أن جميع الحدود، باستثناء الخطأ العشوائي، يمكن أن توضع في الجانب الأيسر من المعادلة. ولذا فإن الخطأ العشوائي يمكن عزله على الجانب الأيمن. على سبيل المثال يمكن التعبير عن (8A.3a) على النحو :

$$\log(Y_{1t}) - a_0 - a_1 \left(\frac{Y_{2t}}{1 + b_2 X_{2t}} \right) - a_2 \left(\frac{X_{1t}}{Y_{2t}} \right) = \varepsilon_{1t}, \quad (8A.4)$$

وبعمومية أكثر دع Y_{mt}, \dots, Y_{1t} المتغيرات الداخلية للنموذج، X_{pt}, \dots, X_{1t}

* إذا كان ينبغي تقدير كل معادلة في النموذج، يتطلب تحليلنا حينئذ أن يعبر عن كل معادلة بالشكل (8A.5).

المتغيرات الخارجية و u_{1t}, \dots, u_{mt} الأخطاء العشوائية. سنفترض، بعدئذ، أن المعادلة التي نرغب في تقديرها مثلاً (رقم i) يمكن التعبير عنها في الشكل ^{*}:

$$F_i(Y_{1t}, \dots, Y_{mt}, X_{1t}, \dots, X_{pt}) = u_{it}, \quad (8A.5)$$

حيث إن الجانب الأيسر من (8A.5) دالة في واحد أو أكثر من المتغيرات Y_{mt}, \dots, Y_{1t} و X_{pt}, \dots, X_{1t} الذي يحتوي على معلومات غير معلومة. وعلى سبيل التوضيح، ففي المعادلة (8A.3a) مثلاً، ستكون هذه الدالة الجانب الأيسر في (8A.4).

إن افتراض إمكانية التعبير عن المعادلة بدلالة الخطأ العشوائي القابل للإضافة، ومن ثم، تصبح في شكل مماثل لـ (8A.5) ليس افتراضاً مقيداً جداً، نظراً لطبيعة النماذج التي يهتم بها الاقتصاديون عادة. على سبيل المثال، يتطلب هذا الافتراض إمكانية حل المعادلة المعينة من النموذج موضع الاهتمام للحصول على الخطأ العشوائي، وعلى سبيل المثال، إذا كانت المعادلة الأولى من النموذج تأخذ الشكل التالي:

$$Y_{1t} = a_0 X_{1t}^{a_1} Y_{2t}^{a_2} e^{u_{1t}}, \quad (8A.6)$$

فإن الشكل المناظر لـ (8A.5) هو:

$$\log(Y_{1t}) - \log(a_0) - a_1 \log(X_{1t}) - a_2 \log(Y_{2t}) = u_{1t}. \quad (8A.7)$$

وعلى سبيل توضيح آخر، اعتبر معادلة من الشكل:

$$Y_{1t} = a_0 + a_1 \left(\frac{e^{a_2 Y_{2t}}}{1 + u_{1t}} \right) + a_3 X_{1t}, \quad (8A.8)$$

حيث إن مدى القيم الممكنة للخطأ العشوائي u_{1t} يكون على النحو $1 + u_{1t} > 0$ ، حينئذ، فإنه يمكن الحصول على الشكل المناظر لـ (8A.5) بملاحظة أولاً أن:

$$\left(\frac{Y_{1t} - a_0 - a_3 X_{1t}}{a_1 e^{a_2 Y_{2t}}} \right) = \frac{1}{1 + u_{1t}}, \quad (8A.9)$$

ولذا، فإن:

$$\left(\frac{a_1 e^{a_2 Y_{2t}}}{Y_{1t} - a_0 - a_3 X_{1t}} \right) - 1 = u_{1t}. \quad (8A.10)$$

وقبل الدخول في قضية التقدير، ومن ثم اختبار الفروض، نعرض هنا نتيجة أولية.

نتيجة أولية

اعتبر نموذج الانحدار العادي المكون من معادلة واحدة :

$$Y_t = b_0 + b_1 X_{1t} + \dots + b_k X_{kt} + u_t, \quad t=1, \dots, N, \quad (8A.11)$$

حيث لاتعاني المتغيرات المستقلة الارتباط الخطي المتعدد، كما أن الخطأ العشوائي يحقق افتراضاتنا المعتادة كافة، وبالتحديد، نفترض أن u_t مستقل عن المتغيرات الداخلية المبطة والحالية والقيم المستقبلية لها كافة. وله قيمة متوقعة صفرية $E(u_t) = 0$ وتباين ثابت $E(u_t^2) = \sigma_u^2$ ، كما أنه ليس مرتبطاً ذاتياً.

تذكر أن الافتراض الأساسي بشأن معاملات الانحدار b_0, b_1, \dots, b_k في نموذج مثل (8A.11) هو أن هذه المعاملات ثوابت، أي أن قيمها لاتعتمد على t . لذا فإن بعض هذه المعاملات أو كلها من الممكن أن يكون صفرًا. ومن الواضح أنه إذا كانت جميع هذه المعاملات تساوي الصفر، يكون المتغير التابع $Y_t = u_t$ ، لذلك تنخفض Y_t إلى متغير عشوائي مستقل عن المتغيرات المستقلة في النموذج.

دع \hat{b}_i مقدار b_i حصل عليه بطريقة المتغير المساعد المعادلة لطريقة المربعات الصغرى. في هذه الحال، فإن \hat{b}_i مقدار غير متحيز لـ b_i (كما أوضحنا في الفصل الرابع) أي $E(\hat{b}_i) = b_i$. وتظل هذه النتيجة صحيحة بغض النظر عما إذا كانت b_i تساوي الصفر أم لا. وقد أوضحنا، أيضاً، في ملحق الفصل الرابع أن تباين \hat{b}_i يمكن التعبير عنه على النحو التالي :

$$\text{var}(\hat{b}_i) = \frac{\sigma_u^2}{N \sum_{t=1}^N \hat{v}_{it}^2} \quad (8A.12)$$

حيث إن \hat{v}_{it} هو المتبقي من انحدار X_{it} على المتغيرات المستقلة الأخرى كافة (مشملة على الحد الثابت في (8A.11)).

ومقام (8A.12) هو مجموع حدود عددها N ، وكل من هذه الحدود، إما أكبر من الصفر أو يساويه. لذلك (ومن بين أشياء أخرى)، فإن قيمة تباين \hat{b}_i تعتمد على حجم العينة N ، ويمكن إثبات أنه (في ظل افتراضات فنية إضافية) إذا كانت N لانهائية فإن المقام في (8A.12) سيكون لانهائياً ولذا، سيكون تباين \hat{b}_i مساوياً للصفر. اعتبر الآن الحالة التي تكون فيها معاملات الانحدار كافة في (8A.11) مساوية للصفر $(b_i = 0, i=0, \dots, k)$. تعني نتائجنا، في هذه الحال، أن $E(\hat{b}_i) = 0, i = 0, \dots, k$ وأنه إذا كان $\text{var}(\hat{b}_i) = 0: N = \infty$ و $i = 0, \dots, k$. فبديهياً، إذا كانت معاملات الانحدار صفراً وحجم العينة لانهائياً فإن القيمة المتوقعة لكل معلمة من معاملات الانحدار ستساوي الصفر، كما أن توقع مربع انحراف القيم عن الصفر (تباينها) سيكون، أيضاً، صفراً. ولذا نتوقع أن قيمة كل مقدرات معاملات الانحدار تساوي الصفر في هذه الحالة. ولكن، يمكن التعبير عن ذلك تعبيراً اصطلاحياً بالقول إنه إذا كانت $E(\hat{b}_i) = 0, i = 0, \dots, k$ وأن $\lim_{N \rightarrow \infty} \text{var}(\hat{b}_i) = 0$ ، حيث، فإن $\text{prob}(|\hat{b}_i - 0| > \delta) = \lim_{N \rightarrow \infty} \text{prob}(|\hat{b}_i| > \delta) = 0$ حيث إن δ هي رقم ثابت موجب صغير*. ونعبر عن ذلك لفظياً بالقول أنه، مع تحقيق هذه الشروط، فإن احتمال أن يكون \hat{b}_i مختلفاً عن الصفر بأي مقدار موجب صغير هو الصفر (أي \hat{b}_i تؤول في الاحتمال إلى الصفر).

والنتيجة المهمة من وراء كل ذلك هي أنه إذا انحدر أحد المتغيرات ذا القيمة المتوسطة الصفرية، والتباين الثابت وغير المرتبط ذاتياً (مثل الخطأ العشوائي) بمجموعة من المتغيرات المستقلة فإن مقدرات معاملات الانحدار سوف تؤول في الاحتمال إلى الصفر، في حال تحقق افتراضات فنية (ومعقولة) إضافية، لذلك فإن القيمة المحسوبة للتعبير، على سبيل المثال، في الحال السابقة:

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_{1i} + \dots + \hat{b}_k X_{ki}, \quad (8A.13)$$

* النص الإنجليزي لهذه العبارة هو:

if $E(\hat{b}_i) = 0, i = 0, \dots, k$, and $\lim_{N \rightarrow \infty} \text{var}(\hat{b}_i) = 0$, Then

$\lim_{N \rightarrow \infty} \text{prob}(|\hat{b}_i - 0| > \delta) = \lim_{N \rightarrow \infty} \text{prob}(|\hat{b}_i| > \delta) = 0$, where δ is positive constant.

سوف يؤول، أيضاً، في الاحتمال إلى الصفر عندما تؤول N إلى ∞ .

طريقة التقدير

سنوضح في هذا البحث، بداية، طريقة تقدير معادلة غير خطية في كل من المتغيرات الداخلية والمعلومات بالمعادلة (8A.3b)، وبعدئذ، سنعمم نتائجنا. افترض أنه يتوافر لدينا عدد N مشاهدات عن المتغيرات في النموذج المكون من معادلتين في (8A.3) وهي Y_{1t} ، Y_{2t} ، X_{1t} ، X_{2t} و X_{3t} . افترض أن المتغيرات X_{3t} و X_{2t} ، X_{1t} ليست مرتبطة خطياً ببعضها بعضاً. افترض، أيضاً، أن الأخطاء العشوائية كافة لها قيم متوسطة صفرية وتباين ثابت، وتغاير ثابت، ولا يعاني أي منها الارتباط الذاتي، وأن كلا منها مستقل عن القيم المبطة والحالية والمستقبلية للمتغيرات الثلاثة الخارجية. وتعتمد خاصية الاتساق لطريقة التقدير التي سنضعها بوضعها على بعض الافتراضات الفنية الإضافية نفترض أنها سوف تتحقق في الواقع. ولأن عرض هذه الافتراضات الفنية وفهمها يتطلب أدوات رياضية وإحصائية أعلى من مستوى هذا الكتاب، سنفترض، ببساطة، تحقق هذه الشروط دون أن نحددها فعلياً.

يمكن التعبير عن المعادلة (8A.3b) في شكل المعادلة (8A.5) على النحو :

$$(Y_{2t}^\alpha X_{3t}) - b_0 - b_1 Y_{1t} - b_2 Y_{1t}^2 - b_3 X_{2t} = \varepsilon_{2t}. \quad (8A.14)$$

ونرمز إلى الجانب الأيسر من (8A.14) بالرمز F_t . افترض أننا اخترنا قيمة افتراضية لكل معلمة تظهر في الجانب الأيسر من (8A.14)، وقمنا ببناء مشاهدات «تقديرات» F_t (مثلاً، F_t^a) طبقاً لهذه القيم. افترض، على سبيل المثال، أننا اخترنا 0.1 لـ α و 3.7 لـ b_0 ، -1.5 لـ b_1 ، 20 لـ b_2 ، وأخيراً -10 لـ b_3 حيثئذ، تحدد مشاهداتنا التقريبية على النحو :

$$F_t^a = (Y_{2t}^{0.1} X_{3t}) - 3.7 + 1.5 Y_{1t} - 2 Y_{1t}^2 + 10 X_{2t}. \quad (8A.15)$$

لاحظ، من (8A.14)، أنه إذا كانت القيم الحقيقية للمعلومات تعادل قيمنا المختارة فإن $F_t^a = F_t = \varepsilon_{2t}$ لجميع قيم t . ومن ناحية أخرى، إذا كانت واحدة أو أكثر من القيم المختارة لاتساوي القيم الحقيقية المناظرة، فإن $F_t^a \neq \varepsilon_{2t}$ لجميع قيم t .

افترض، للحظة، أن قيم المعلومات التي اخترناها هي القيم الحقيقية. لذلك، فإن $\hat{F}_t^a = F_t = \varepsilon_{2t}$ افترض، أيضاً، أننا أجرينا، ماسنطلق عليه من الآن فصاعداً، انحدار المرحلة الأولى للمتغير F_t المبني على المتغيرات الخارجية للنموذج ومربعاتها، أي $X_{1t}, X_{2t}, X_{3t}, X_{1t}^2, X_{2t}^2, X_{3t}^2$ ، وعلى الحد الثابت وسنناقش أدناه، القضية المرتبطة باختيار المتغيرات المستقلة للمرحلة الأولى. وبالعودة إلى تحليلنا، دع \hat{F}_t ترمز إلى القيمة المحسوبة لـ F_t من انحدار المرحلة الأولى، حينئذ، تتضمن النتائج الأولية المعطاة في المبحث السابق مباشرة أنه طالما أن $F_t = \varepsilon_{2t}$ هو خطأ عشوائي يحقق افتراضاتنا المعتادة كافة، وطالما أن F_t مستقل عن جميع المتغيرات الخارجية الثلاثة فإن \hat{F}_t سوف تؤول في الاحتمال إلى الصفر عندما تقارب N من ∞ . اعتبر التجميع:

$$S = \sum_{t=1}^N \frac{\hat{F}_t^2}{N}. \quad (8A.16)$$

حينئذ ينبغي أن يكون واضحاً أن S سوف تؤول، أيضاً، في الاحتمال إلى الصفر عندما تؤول N إلى ∞ .

افترض الآن أن قيمنا المختارة للمعلومات ليست هي القيم الحقيقية، حينئذ، (وكما اشرنا) فإن متغيرنا المكون $F_t^a \neq \varepsilon_{2t}$, $t=1, \dots, N$. في هذه الحالة، تكون قيمة F_t^a دالة في متغيرات المعادلة أي في $X_{1t}, Y_{1t}, X_{2t}, Y_{2t}, X_{3t}$. افترض الآن، كما في الحالة السابقة أنه تم انحدار F_t^a على المتغيرات المستقلة في المرحلة الأولى $X_{1t}, X_{2t}, X_{3t}, X_{1t}^2, X_{2t}^2, X_{3t}^2$ ، والحد الثابت. دع القيم المحسوبة الناتجة لـ F_t^a هي \hat{F}_t^a . نتوقع، حينئذ، أن متوسط مجموع المربعات للقيم المحسوبة أي أن:

$$S^a = \sum_{t=1}^N \frac{(\hat{F}_t^a)^2}{N}, \quad (8A.17)$$

لن تؤول في الاحتمال إلى الصفر. هذه هي الحالة بالضبط، مع وجود الافتراضات الفنية الإضافية. ولأن S^a هو مجموع مربعات، فإنه يمكن إثبات أن، مع توافر الفروض الفنية الإضافية، S^a سوف تؤول في الاحتمال إلى رقم موجب.

النقطة الأساسية في المناقشة السابقة هي أن متوسط الحدود $(\hat{F}_i^a)^2$ سوف يؤول إلى الصفر إذا كانت القيم المختارة للمعلمات هي القيم الحقيقية، وسوف يؤول إلى رقم موجب في الحالات الأخرى. وهكذا، فإن طريقة التقدير التالية واضحة. ابحث عن القيم الممكنة لمعلمات الانحدار حتى تجد المجموعة التي تصغر متوسط مربعات المتغيرات المحسوبة $(\hat{F}_i^a)^2$. نأخذ هذه المجموعة من القيم كتقديراتنا لمعلمات الانحدار. هذه المقدرات متسقة لأن حجم العينة لا نهائي، ويصل متوسط مربعات المتغيرات المحسوبة، $(\hat{F}_i^a)^2$ إلى حده الأدنى (عند الصفر) بوساطة القيم الحقيقية للمعلمات.

اختيار المتغيرات المستقلة للمرحلة الأولى

تشابه قضايا اختيار المتغيرات المستقلة للمرحلة الأولى، إلى حد كبير، تلك التي ناقشناها في المبحث (٨-٣). على سبيل المثال، إذا كان حجم العينة لانهائياً، فإن تباينات مقدرات معلمات النموذج ستكون مرتبطة عكسياً في انحدار المرحلة الأولى بقوى المتغيرات الخارجية، لذلك، فإن قوى أعلى وأعلى (مع القوى الأقل) لهذه المتغيرات ينبغي أن تؤخذ في الحسبان في انحدار المرحلة الأولى. ولكن، عند التطبيق، يكون حجم العينة محدوداً، ولذلك، فإن عدد الحدود المستخدمة في انحدار المرحلة الأولى ينبغي أن يكون محدوداً. وفي الحقيقة، يمكن إثبات أن استخدام عدد كبير من الحدود في انحدار المرحلة الأولى (مثلاً P يساوي حجم العينة N) يؤدي إلى الحصول على مقدرات من نوع مقدرات المربعات الصغرى، وهذه المقدرات غير متسقة.* وكما في المبحث (٨-٣) تكون لدينا

* للقراء الأكثر دراية، إذا كانت $P=N$ ، حيث $\hat{F}_i^a = F_i^a$. لذلك، سوف نختار قيم المعلمات لتقليل $\sum_{i=1}^N (\hat{F}_i^a)^2 / N$ إلى حدها الأدنى، والتي تكون معادلة لتدنية مجموع مربعات الأخطاء العشوائية المقررة (أي طريقة المربعات الصغرى غير الخطية). ولكن طرق المربعات الصغرى لا تعطي مقدرات متسقة إذا كان نموذج الانحدار، من بين أشياء أخرى، له متغيرات داخلية في الجانب الأيمن المعادلة.

معضلة، ومرة ثانية، نقترح اختيار P بحيث تكون $N-P \geq 20$. وبالطبع، تتطلب خاصية الاتساق أن تكون $N-P$ لانهائية. نلاحظ، أيضاً، بدون إثبات أن خاصية الاتساق لطريقة التقدير تتطلب أن تكون P ، في الأقل، مماثلة لعدد الملاحظات التي ينبغي تقديرها.

استعراض عام وخطوط عريضة للطريقة

يمكننا الآن أن نعرض عرضاً عاماً للخطوط العريضة لطريقة تقدير المعادلة غير الخطية في كل من المتغيرات الداخلية والملاحظات في الخطوات التالية :

- (١) أكتب المعادلة في شكل المعادلة (8A.5) وضع رمزاً للجانب اليسر منها F_t .
 - (٢) دع F_t^a القيمة التقريبية لـ F_t والتي تحدد عن طريق اختيار مجموعة من القيم للملاحظات التي تظهر في المعادلة.
 - (٣) حدد الأشكال متعددة الحدود للمتغيرات الخارجية التي تحدد المتغيرات المستقلة للمرحلة الأولى، تذكر أن عدد المتغيرات المستقلة في المرحلة الأولى، مثلاً P ، ينبغي أن يكون، في الأقل، كعدد الملاحظات التي تظهر في المعادلة التي ينبغي تقديرها. وطالما أن حجم العينة في التطبيق العملي يكون محدوداً N . اختر P بحيث تكون $(N-P)$ تعادل، في الأقل، 20.
 - (٤) دع \hat{F}_t^a القيمة المحسوبة لـ F_t^a والتي يحصل عليها من انحدار F_t^a على المتغيرات المستقلة للمرحلة الأولى.
 - (٥) ابحث عن القيم الممكنة للملاحظات لإيجاد مجموعة من القيم التي تدني $\sum_{t=1}^N (\hat{F}_t^a)^2 / N$ واعتبر هذه القيم تقديرات للملاحظات المناظرة.
- من الواضح أن التنفيذ العملي لهذه الطريقة سيتطلب معلومات في التحليل العددي والبرمجة بالحاسوب، أو إتاحة برنامج حاسوب متخصص يحتوي على هذا المنهج باعتباره خياراً.

اختبار الفرضيات، فترات الثقة وتباينات العينة الكبيرة: تعليق

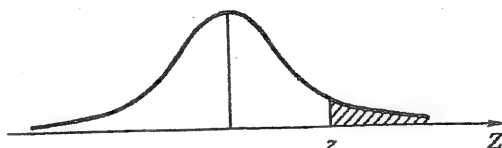
لسوء الحظ، لا يمكننا، من الآن فصاعداً، أن نعطي قراءنا الأكثر دراية، صيغاً لتباينات العينة الكبيرة للمقدرات وذلك لأن عرضها يتطلب أدوات رياضية تتجاوز مستوى هذا الكتاب.* ولكن، إذا كان برنامج الحاسوب المستخدم يحتوي على هذا المنهج بوصفه خياراً، فسيطبع (من بين أشياء أخرى) تقديرات المعلومات وتقديرات التباينات للعينات الكبيرة المناظرة. ويمكن إثبات أنه، مع تحقيق فروض فنية إضافية، إذا كان حجم العينة لا نهائياً، فإن مقدرات المعلومات سوف تكون موزعة توزيعاً طبيعياً. لذلك، يمكن اختبار الفرضيات المرتبطة بقيمة إحدى المعلومات الفردية أو بناء فترات الثقة، بدلالة ناتج مثل هذا البرنامج، عن طريق استخدام التوزيع الطبيعي على سبيل التقريب. افترض، مثلاً، أن a_2 هي واحدة من المعلومات، والناتج الذي يطبعه الحاسوب هو $\hat{a}=10$ و $\hat{\sigma}_{a_2}^2=16$. حيثئذ، فإن 95% فترة ثقة لـ a_2 ، المبنية على التوزيع الطبيعي للعينة الكبيرة ستكون $10 \pm 4(1.96)$ أو 10 ± 7.84 .

* بالنسبة للقراء الأكثر دراية، افترض أن المعادلة التي ينبغي تقديرها تأخذ شكل رقم (8A.5)، ونرمز للجانب الأيسر للرمز F_{ii} . افترض أن المعادلة تحتوي على معاملات a_0, \dots, a_k . دع $f_{ii} = (\partial F_i / \partial a_i)$, $i = 0, \dots, k$. سوف تتضمن F_{ii} عموماً واحداً أو أكثر من المعلمات a_0, \dots, a_k . احصل على عدد N مشاهدات عن كل من f_{ii} عن طريق التعويض عن المعلمات المتضمنة بمقدرات متسقة لها. والآن، دع \hat{f}_{ii} . ترمز إلى القيمة المحسوبة لـ f_{ii} من انحدار f_{ii} على المتغيرات المستقلة في المرحلة الأولى كافة. دع أخيراً \hat{f}_{ii} المتبقي رقم t من انحدار \hat{f}_{ii} على k متغيرات \hat{f}_{ii} ، حيث إن \hat{f}_{ii} ، حيث \hat{f}_{ii} ، يقدر تباين العينة الكبيرة لـ \hat{a}_i تقديراً متسقاً على النحو $\hat{\sigma}^2 / \sum_{i=1}^N \hat{f}_{ii}^2$ حيث إن $\hat{\sigma}^2$ هي مقدر متسق لتباين الخطأ العشوائي للمعادلة.

الجدول الإحصائية

جدول (١). التوزيع الطبيعي المعياري

$$Z = \frac{X - \mu}{\sigma}$$

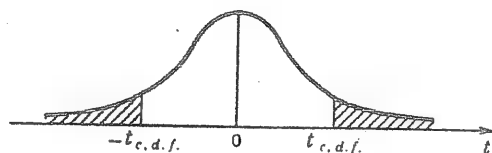


z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010

The table plots the cumulative probability $Z \geq z$.

SOURCE: Reprinted from Edward J. Kane, *Economic Statistics and Econometrics: An Introduction to Quantitative Economics*, New York: Harper & Row, Publishers, 1968.

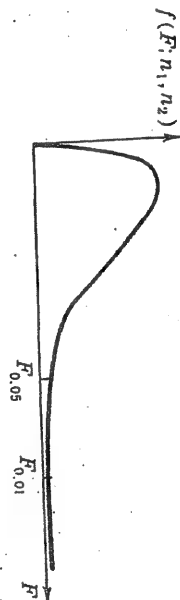
جدول (٢). التوزيع - t



Degrees of Freedom	Probability of a Value Greater in Absolute Value than the Table Entry					
	0.01	0.02	0.05	0.1	0.2	0.3
1	63.657	31.821	12.706	6.314	3.078	1.963
2	9.925	6.965	4.303	2.920	1.886	1.386
3	5.841	4.541	3.182	2.353	1.638	1.250
4	4.604	3.747	2.776	2.132	1.533	1.190
5	4.032	3.365	2.571	2.015	1.476	1.156
6	3.707	3.143	2.447	1.943	1.440	1.134
7	3.499	2.998	2.365	1.895	1.415	1.119
8	3.355	2.896	2.306	1.860	1.397	1.108
9	3.250	2.821	2.262	1.833	1.383	1.100
10	3.169	2.764	2.228	1.812	1.372	1.093
11	3.106	2.718	2.201	1.796	1.363	1.088
12	3.055	2.681	2.179	1.782	1.356	1.083
13	3.012	2.650	2.160	1.771	1.350	1.079
14	2.977	2.624	2.145	1.761	1.345	1.076
15	2.947	2.602	2.131	1.753	1.341	1.074
16	2.921	2.583	2.120	1.746	1.337	1.071
17	2.898	2.567	2.110	1.740	1.333	1.069
18	2.878	2.552	2.101	1.734	1.330	1.067
19	2.861	2.539	2.093	1.729	1.328	1.066
20	2.845	2.528	2.086	1.725	1.325	1.064
21	2.831	2.518	2.080	1.721	1.323	1.063
22	2.819	2.508	2.074	1.717	1.321	1.061
23	2.807	2.500	2.069	1.714	1.319	1.060
24	2.797	2.492	2.064	1.711	1.318	1.059
25	2.787	2.485	2.060	1.708	1.316	1.058
26	2.779	2.479	2.056	1.706	1.315	1.058
27	2.771	2.473	2.052	1.703	1.314	1.057
28	2.763	2.467	2.048	1.701	1.313	1.056
29	2.756	2.462	2.045	1.699	1.311	1.055
30	2.750	2.457	2.042	1.697	1.310	1.055
∞	2.576	2.326	1.960	1.645	1.282	1.036

SOURCE: Reprinted from Table IV in Sir Ronald A. Fisher, *Statistical Methods for Research Workers*, 13th edition. Oliver & Boyd Ltd., Edinburgh, 1963, with the permission of the publisher and the late Sir Ronald Fisher's Literary Executor.

جدول (٣). القيم الحرجة لتوزيع F-

n₁ degrees of freedom (for greater mean square)

n ₂	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞	n ₂
1	161	200	216	225	230	234	237	239	241	242	243	244	245	246	248	249	250	251	252	253	254	254	254	254	1
2	4.052	4.999	5.403	5.625	5.764	5.859	5.928	5.981	6.022	6.056	6.082	6.106	6.142	6.169	6.208	6.234	6.258	6.286	6.302	6.313	6.314	6.314	6.314	6.314	2
3	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41	19.42	19.43	19.44	19.45	19.46	19.47	19.48	19.49	19.49	19.49	19.50	19.50	3
4	98.49	99.00	99.17	99.25	99.30	99.33	99.34	99.36	99.38	99.40	99.41	99.42	99.43	99.44	99.45	99.46	99.47	99.48	99.48	99.49	99.49	99.49	99.50	99.50	4
5	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.76	8.74	8.71	8.69	8.66	8.64	8.62	8.60	8.58	8.57	8.56	8.54	8.54	8.53	5
6	14.12	13.02	12.46	12.11	11.91	11.79	11.71	11.65	11.60	11.56	11.53	11.51	11.49	11.47	11.45	11.43	11.41	11.39	11.37	11.35	11.34	11.33	11.32	11.31	6
7	17.20	16.00	15.09	14.38	13.82	13.44	13.16	12.94	12.77	12.64	12.54	12.46	12.39	12.33	12.28	12.24	12.20	12.17	12.14	12.12	12.11	12.10	12.09	12.08	7
8	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.41	4.40	4.39	4.38	8
9	16.26	13.27	12.06	11.39	10.97	10.67	10.45	10.27	10.15	10.05	9.96	9.89	9.77	9.68	9.55	9.47	9.38	9.29	9.24	9.17	9.13	9.07	9.04	9.02	9
10	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.72	3.71	3.69	3.68	3.67	10
11	13.74	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72	7.60	7.52	7.39	7.31	7.23	7.14	7.09	7.02	6.99	6.94	6.90	6.88	11
12	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57	3.52	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.28	3.25	3.24	3.23	12
13	12.25	9.55	8.45	7.85	7.46	7.19	7.00	6.84	6.71	6.62	6.54	6.47	6.35	6.27	6.15	6.07	5.98	5.90	5.85	5.78	5.75	5.70	5.67	5.65	13
14	5.13	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28	3.23	3.20	3.15	3.12	3.08	3.05	3.03	3.00	2.98	2.96	2.94	2.93	14
15	11.26	8.65	7.59	7.01	6.63	6.37	6.19	6.03	5.91	5.82	5.74	5.67	5.56	5.48	5.36	5.28	5.20	5.11	5.06	5.00	4.96	4.91	4.88	4.86	15
16	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07	3.02	2.98	2.93	2.90	2.86	2.82	2.80	2.77	2.76	2.73	2.72	2.71	16
17	10.56	8.02	6.99	6.42	6.06	5.80	5.62	5.47	5.35	5.26	5.18	5.11	5.00	4.92	4.80	4.73	4.64	4.56	4.51	4.45	4.41	4.36	4.33	4.31	17
18	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91	2.86	2.82	2.77	2.74	2.70	2.67	2.64	2.61	2.59	2.56	2.55	2.54	18
19	10.04	7.56	6.55	5.99	5.64	5.39	5.21	5.06	4.95	4.85	4.78	4.71	4.60	4.52	4.41	4.33	4.25	4.17	4.12	4.05	4.01	3.96	3.93	3.91	19
20	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.83	2.79	2.74	2.70	2.65	2.61	2.57	2.53	2.50	2.47	2.45	2.42	2.41	2.40	20
21	9.65	7.20	6.22	5.67	5.32	5.07	4.88	4.74	4.63	4.54	4.46	4.40	4.29	4.21	4.10	4.02	3.94	3.86	3.80	3.74	3.70	3.66	3.62	3.60	21
22	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69	2.64	2.60	2.54	2.50	2.46	2.42	2.40	2.36	2.35	2.32	2.31	2.30	22
23	9.33	6.93	5.95	5.41	5.06	4.82	4.65	4.50	4.39	4.30	4.22	4.16	4.05	3.98	3.86	3.78	3.70	3.61	3.56	3.49	3.46	3.41	3.38	3.36	23
24	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.63	2.60	2.55	2.51	2.46	2.42	2.38	2.34	2.32	2.28	2.26	2.24	2.22	2.21	24
25	9.07	6.70	5.74	5.20	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96	3.85	3.78	3.67	3.59	3.51	3.42	3.37	3.30	3.27	3.21	3.18	3.16	25

تابع جدول (3)

n_2	n_1 degrees of freedom (for greater mean square)																										n_1
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞			
14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.56	2.53	2.48	2.44	2.39	2.35	2.31	2.27	2.24	2.21	2.19	2.16	2.14	2.13	14		
15	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55	2.51	2.48	2.43	2.39	2.33	2.29	2.25	2.21	2.18	2.15	2.12	2.10	2.08	2.07	15		
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.42	2.37	2.33	2.28	2.24	2.20	2.16	2.13	2.09	2.07	2.04	2.02	2.01	16		
17	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45	2.41	2.38	2.33	2.29	2.23	2.19	2.15	2.11	2.08	2.04	2.02	1.99	1.97	1.96	17		
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.29	2.25	2.19	2.15	2.11	2.07	2.04	2.00	1.98	1.95	1.93	1.92	18		
19	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43	2.38	2.34	2.31	2.26	2.21	2.15	2.11	2.07	2.02	2.00	1.96	1.94	1.91	1.88	1.87	19		
20	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40	2.35	2.31	2.28	2.23	2.18	2.12	2.08	2.04	1.99	1.96	1.92	1.90	1.87	1.85	1.84	20		
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.20	2.15	2.09	2.05	2.00	1.96	1.93	1.89	1.87	1.84	1.82	1.81	21		
22	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35	2.30	2.26	2.23	2.18	2.13	2.07	2.03	1.98	1.93	1.91	1.87	1.84	1.81	1.80	1.78	22		
23	4.28	3.42	3.03	2.80	2.64	2.53	2.45	2.38	2.33	2.28	2.24	2.21	2.16	2.10	2.04	2.00	1.96	1.91	1.88	1.84	1.82	1.79	1.77	1.76	23		
24	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30	2.26	2.22	2.18	2.13	2.09	2.02	1.98	1.94	1.89	1.86	1.82	1.80	1.76	1.74	1.73	24		
25	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24	2.20	2.16	2.11	2.06	2.00	1.96	1.92	1.87	1.84	1.80	1.77	1.74	1.72	1.71	25		
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.10	2.05	1.99	1.95	1.90	1.85	1.82	1.78	1.76	1.72	1.70	1.69	26		
	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.17	3.09	3.02	2.96	2.86	2.77	2.66	2.58	2.50	2.41	2.36	2.28	2.25	2.19	2.15	2.13			

الجدول الإحصائي

مقدمة في الاقتصاد القياسي

تابع جدول (٣)

n_1	n_2 degrees of freedom (for greater mean square)																								n_2
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	x	
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.30	2.25	2.20	2.16	2.13	2.08	2.03	1.97	1.93	1.88	1.84	1.80	1.76	1.74	1.71	1.68	1.67	27
28	4.20	3.34	2.95	2.71	2.56	2.44	2.36	2.29	2.24	2.19	2.15	2.12	2.06	2.02	1.96	1.91	1.87	1.81	1.78	1.75	1.72	1.69	1.67	1.65	28
29	4.18	3.33	2.93	2.70	2.54	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.05	2.00	1.94	1.90	1.85	1.80	1.77	1.73	1.71	1.68	1.65	1.64	29
30	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.21	2.16	2.12	2.09	2.04	1.99	1.93	1.89	1.84	1.79	1.76	1.72	1.69	1.66	1.64	1.62	30
32	4.15	3.30	2.90	2.67	2.51	2.40	2.32	2.25	2.19	2.14	2.10	2.07	2.02	1.97	1.91	1.86	1.82	1.76	1.74	1.69	1.67	1.64	1.61	1.59	32
34	4.13	3.28	2.88	2.65	2.49	2.38	2.30	2.23	2.17	2.12	2.08	2.05	2.00	1.95	1.89	1.84	1.80	1.74	1.71	1.67	1.64	1.61	1.59	1.57	34
36	4.11	3.26	2.86	2.63	2.48	2.36	2.28	2.21	2.15	2.10	2.06	2.03	1.98	1.93	1.87	1.82	1.78	1.72	1.69	1.65	1.62	1.59	1.56	1.55	36
38	4.10	3.25	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02	1.96	1.92	1.85	1.80	1.76	1.71	1.67	1.63	1.60	1.57	1.54	1.53	38
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.07	2.04	2.00	1.95	1.90	1.84	1.79	1.74	1.69	1.66	1.61	1.59	1.55	1.51	1.51	40
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	2.02	1.99	1.94	1.89	1.82	1.78	1.73	1.68	1.64	1.60	1.57	1.54	1.51	1.49	42
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	2.01	1.98	1.93	1.88	1.81	1.76	1.72	1.66	1.63	1.58	1.56	1.52	1.50	1.48	44
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.14	2.09	2.04	2.00	1.97	1.92	1.87	1.80	1.75	1.71	1.65	1.62	1.57	1.54	1.51	1.48	1.46	46
48	4.04	3.19	2.80	2.56	2.41	2.30	2.21	2.14	2.08	2.03	1.99	1.96	1.91	1.86	1.79	1.74	1.70	1.64	1.61	1.56	1.53	1.50	1.47	1.45	48
	7.19	5.08	4.22	3.74	3.42	3.20	3.04	2.90	2.80	2.71	2.64	2.58	2.48	2.40	2.28	2.20	2.11	2.02	1.96	1.88	1.84	1.78	1.73	1.70	

(٣) تابع جدول

n_2	n_1 degrees of freedom (for greater mean square)																							
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.02	1.98	1.90	1.85	1.78	1.74	1.69	1.63	1.59	1.56	1.50	1.48	1.46	1.44	1.44
55	4.02	3.17	2.78	2.54	2.38	2.27	2.18	2.11	2.05	2.00	1.97	1.93	1.88	1.83	1.76	1.72	1.67	1.61	1.58	1.52	1.50	1.46	1.43	1.41
60	4.00	3.15	2.76	2.52	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.86	1.81	1.75	1.70	1.65	1.59	1.56	1.50	1.48	1.44	1.41	1.39
65	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.02	1.98	1.94	1.90	1.85	1.80	1.73	1.68	1.63	1.57	1.54	1.49	1.46	1.44	1.41	1.39
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.01	1.97	1.93	1.89	1.84	1.79	1.72	1.67	1.62	1.56	1.53	1.47	1.45	1.40	1.37	1.35
75	3.97	3.12	2.73	2.49	2.34	2.22	2.13	2.06	2.00	1.96	1.92	1.88	1.83	1.78	1.71	1.66	1.61	1.55	1.52	1.46	1.44	1.40	1.37	1.35
80	3.96	3.11	2.72	2.48	2.33	2.21	2.12	2.05	1.99	1.95	1.91	1.88	1.82	1.77	1.70	1.65	1.60	1.54	1.51	1.45	1.42	1.38	1.35	1.32
85	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.98	1.94	1.90	1.86	1.81	1.75	1.68	1.63	1.57	1.51	1.48	1.42	1.39	1.34	1.30	1.28
90	3.94	3.09	2.70	2.46	2.30	2.19	2.10	2.03	1.97	1.92	1.88	1.85	1.79	1.75	1.68	1.63	1.57	1.51	1.48	1.42	1.39	1.34	1.30	1.28
95	3.93	3.08	2.69	2.45	2.30	2.19	2.10	2.03	1.97	1.92	1.88	1.85	1.79	1.75	1.68	1.63	1.57	1.51	1.48	1.42	1.39	1.34	1.30	1.28
100	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.95	1.90	1.86	1.83	1.77	1.72	1.65	1.60	1.55	1.49	1.45	1.39	1.36	1.31	1.27	1.25
105	3.91	3.06	2.67	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.85	1.82	1.76	1.71	1.64	1.59	1.54	1.47	1.44	1.37	1.34	1.29	1.25	1.23
110	3.90	3.05	2.66	2.42	2.26	2.15	2.06	1.99	1.93	1.88	1.84	1.81	1.75	1.70	1.63	1.58	1.53	1.47	1.43	1.37	1.33	1.28	1.24	1.22
115	3.89	3.04	2.65	2.41	2.25	2.14	2.05	1.98	1.92	1.87	1.83	1.80	1.74	1.69	1.62	1.57	1.52	1.45	1.42	1.35	1.32	1.26	1.22	1.20
120	3.88	3.03	2.64	2.40	2.24	2.13	2.04	1.97	1.91	1.86	1.82	1.79	1.73	1.68	1.61	1.56	1.51	1.45	1.41	1.35	1.31	1.26	1.22	1.20
125	3.87	3.02	2.63	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.81	1.78	1.72	1.67	1.60	1.55	1.49	1.45	1.39	1.35	1.29	1.24	1.20	1.18
130	3.86	3.01	2.62	2.38	2.22	2.11	2.02	1.95	1.89	1.84	1.80	1.77	1.71	1.66	1.59	1.54	1.47	1.43	1.37	1.33	1.27	1.22	1.18	1.16
135	3.85	3.00	2.61	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.79	1.76	1.70	1.65	1.58	1.53	1.47	1.43	1.37	1.33	1.27	1.22	1.18	1.16
140	3.84	2.99	2.60	2.36	2.20	2.09	2.00	1.93	1.87	1.82	1.78	1.75	1.69	1.64	1.57	1.52	1.46	1.41	1.35	1.31	1.25	1.20	1.16	1.14
145	3.83	2.98	2.59	2.35	2.19	2.08	1.99	1.92	1.86	1.81	1.77	1.74	1.68	1.63	1.56	1.51	1.45	1.40	1.34	1.30	1.24	1.19	1.15	1.13
150	3.82	2.97	2.58	2.34	2.18	2.07	1.98	1.91	1.85	1.80	1.76	1.73	1.67	1.62	1.55	1.50	1.44	1.39	1.33	1.28	1.22	1.17	1.13	1.11
155	3.81	2.96	2.57	2.33	2.17	2.06	1.97	1.90	1.84	1.79	1.75	1.72	1.66	1.61	1.54	1.49	1.43	1.38	1.32	1.26	1.21	1.16	1.12	1.10
160	3.80	2.95	2.56	2.32	2.16	2.05	1.96	1.89	1.83	1.78	1.74	1.71	1.65	1.60	1.53	1.48	1.42	1.37	1.31	1.25	1.20	1.15	1.11	1.09
165	3.79	2.94	2.55	2.31	2.15	2.04	1.95	1.88	1.82	1.77	1.73	1.70	1.64	1.59	1.52	1.47	1.41	1.36	1.30	1.24	1.19	1.14	1.10	1.08
170	3.78	2.93	2.54	2.30	2.14	2.03	1.94	1.87	1.81	1.76	1.72	1.69	1.63	1.58	1.51	1.46	1.40	1.35	1.29	1.23	1.18	1.13	1.09	1.07
175	3.77	2.92	2.53	2.29	2.13	2.02	1.93	1.86	1.80	1.75	1.71	1.68	1.62	1.57	1.50	1.45	1.39	1.34	1.28	1.22	1.17	1.12	1.08	1.06
180	3.76	2.91	2.52	2.28	2.12	2.01	1.92	1.85	1.79	1.74	1.70	1.67	1.61	1.56	1.49	1.44	1.38	1.33	1.27	1.21	1.16	1.11	1.07	1.05
185	3.75	2.90	2.51	2.27	2.11	2.00	1.91	1.84	1.78	1.73	1.69	1.66	1.60	1.55	1.48	1.43	1.37	1.32	1.26	1.20	1.15	1.10	1.06	1.04
190	3.74	2.89	2.50	2.26	2.10	1.99	1.90	1.83	1.77	1.72	1.68	1.65	1.59	1.54	1.47	1.42	1.36	1.31	1.25	1.19	1.14	1.09	1.05	1.03
195	3.73	2.88	2.49	2.25	2.09	1.98	1.89	1.82	1.76	1.71	1.67	1.64	1.58	1.53	1.46	1.41	1.35	1.30	1.24	1.18	1.13	1.08	1.04	1.02
200	3.72	2.87	2.48	2.24	2.08	1.97	1.88	1.81	1.75	1.70	1.66	1.63	1.57	1.52	1.45	1.40	1.34	1.29	1.23	1.17	1.12	1.07	1.03	1.01
205	3.71	2.86	2.47	2.23	2.07	1.96	1.87	1.80	1.74	1.69	1.65	1.62	1.56	1.51	1.44	1.39	1.33	1.28	1.22	1.16	1.11	1.06	1.02	1.00
210	3.70	2.85	2.46	2.22	2.06	1.95	1.86	1.79	1.73	1.68	1.64	1.61	1.55	1.50	1.43	1.38	1.32	1.26	1.20	1.14	1.09	1.04	1.00	0.98
215	3.69	2.84	2.45	2.21	2.05	1.94	1.85	1.78	1.72	1.67	1.63	1.60	1.54	1.49	1.42	1.37	1.31	1.25	1.19	1.13	1.08	1.03	0.99	0.97
220	3.68	2.83	2.44	2.20	2.04	1.93	1.84	1.77	1.71	1.66	1.62	1.59	1.53	1.48	1.41	1.36	1.30	1.24	1.18	1.12	1.07	1.02	0.98	0.96
225	3.67	2.82	2.43	2.19	2.03	1.92	1.83	1.76	1.70	1.65	1.61	1.58	1.52	1.47	1.40	1.35	1.29	1.23	1.17	1.11	1.06	1.01	0.97	0.95
230	3.66	2.81	2.42	2.18	2.02	1.91	1.82	1.75	1.69	1.64	1.60	1.57	1.51	1.46	1.39	1.34	1.28	1.22	1.16	1.10	1.05	1.00	0.96	0.94
235	3.65	2.80	2.41	2.17	2.01	1.90	1.81	1.74	1.68	1.63	1.59	1.56	1.50	1.45	1.38	1.33	1.27	1.21	1.15	1.09	1.04	0.99	0.95	0.93
240	3.64	2.79	2.40	2.16	2.00	1.89	1.80	1.73	1.67	1.62	1.58	1.55	1.49	1.44	1.37	1.32	1.26	1.20	1.14	1.08	1.03	0.98	0.94	0.92
245	3.63	2.78	2.39	2.15	1.99	1.88	1.79	1.72	1.66	1.61	1.57	1.54	1.48	1.43	1.36	1.31	1.25	1.19	1.13	1.07	1.02	0.97	0.93	0.91
250	3.62	2.77	2.38	2.14	1.98	1.87	1.78	1.71	1.65	1.60	1.56	1.53	1.47	1.42	1.35	1.30	1.24	1.18	1.12	1.06	1.01	0.96	0.92	0.90
255	3.61	2.76	2.37	2.13	1.97	1.86	1.77	1.70	1.64	1.59	1.55	1.52	1.46	1.41	1.34	1.29	1.23	1.17	1.11	1.05	1.00	0.95	0.91	0.89
260	3.60	2.75	2.36	2.12	1.96	1.85	1.76	1.69	1.63	1.58	1.54	1.51	1.45	1.40	1.33	1.28	1.22	1.16	1.10	1.04	0.99	0.94	0.90	0.88
265	3.59	2.74	2.35	2.11	1.95	1.84	1.75	1.68	1.62	1.57	1.53	1.50	1.44	1.39	1.32	1.27	1.21	1.15	1.09	1.03	0.98	0.93	0.89	0.87
270	3.58	2.73	2.34	2.10	1.94	1.83	1.74	1.67	1.61	1.56	1.52	1.49	1.43	1.38	1.31	1.26	1.20	1.14	1.08	1.02	0.97	0.92	0.88	0.86
275	3.57	2.72	2.33	2.09	1.93	1.82	1.73	1.66	1.60	1.55	1.51	1.48	1.42	1.37	1.30	1.25	1.19	1.13	1.07	1.01	0.96	0.91	0.87	0.85
280	3.56	2.71	2.32	2.08	1.92	1.81	1.72	1.65	1.59	1.54	1.50	1.47	1.41	1.36	1.29	1.24	1.18	1.12	1.06	1.00	0.95	0.90	0.86	0.84
285	3.55	2.70	2.31	2.07	1.91	1.80	1.71	1.64	1.58	1.53	1.49	1.46	1.40	1.35	1.28	1.23	1.17	1.11	1.05	0.99	0.94	0.89	0.85	0.83
290	3.54	2.69	2.30	2.06	1.90	1.79	1.70	1.63	1.57	1.52	1.48	1.45	1.39	1.34	1.27	1.22	1.16	1.10	1.04	0.98	0.93	0.88	0.84	0.82
295	3.53	2.68	2.29	2.05	1.89	1.78	1.69	1.62	1.56	1.51	1.47	1.44	1.38	1.33	1.26	1.21	1.15	1.09	1.03	0.97	0.92	0.87	0.83	0.81
300	3.52	2.67	2.28	2.04	1.88	1.77	1.68	1.61	1.55	1.50	1.46	1.43	1.37	1.32	1.25	1.20	1.14	1.08	1.02	0.96	0.91	0.86	0.82	0.80
305	3.51	2.66	2.27	2.03	1.87	1.76	1.67	1.60	1.54	1.49	1.45	1.42	1.36	1.31	1.24	1.19	1.13	1.07	1.01	0.95	0.90	0.85	0.81	

جدول (٤). القيم الحرجة لاختبار ديربان - واتسون

n	k' = 1		k' = 2		k' = 3		k' = 4		k' = 5	
	d _l	d _u	d _l	d _u	d _l	d _u	d _l	d _u	d _l	d _u
15	0.95	1.23	0.83	1.40	0.71	1.61	0.59	1.84	0.48	2.09
16	0.98	1.24	0.86	1.40	0.75	1.59	0.64	1.80	0.53	2.03
17	1.01	1.25	0.90	1.40	0.79	1.58	0.68	1.77	0.57	1.98
18	1.03	1.26	0.93	1.40	0.82	1.56	0.72	1.74	0.62	1.93
19	1.06	1.28	0.96	1.41	0.86	1.55	0.76	1.72	0.66	1.90
20	1.08	1.28	0.99	1.41	0.89	1.55	0.79	1.70	0.70	1.87
21	1.10	1.30	1.01	1.41	0.92	1.54	0.83	1.69	0.73	1.84
22	1.12	1.31	1.04	1.42	0.95	1.54	0.86	1.68	0.77	1.82
23	1.14	1.32	1.06	1.42	0.97	1.54	0.89	1.67	0.80	1.80
24	1.16	1.33	1.08	1.43	1.00	1.54	0.91	1.66	0.83	1.79
25	1.18	1.34	1.10	1.43	1.02	1.54	0.94	1.65	0.86	1.77
26	1.19	1.35	1.12	1.44	1.04	1.54	0.96	1.65	0.88	1.76
27	1.21	1.36	1.13	1.44	1.06	1.54	0.99	1.64	0.91	1.75
28	1.22	1.37	1.15	1.45	1.08	1.54	1.01	1.64	0.93	1.74
29	1.24	1.38	1.17	1.45	1.10	1.54	1.03	1.63	0.96	1.73
30	1.25	1.38	1.18	1.46	1.12	1.54	1.05	1.63	0.98	1.73
31	1.26	1.39	1.20	1.47	1.13	1.55	1.07	1.63	1.00	1.72
32	1.27	1.40	1.21	1.47	1.15	1.55	1.08	1.63	1.02	1.71
33	1.28	1.41	1.22	1.48	1.16	1.55	1.10	1.63	1.04	1.71
34	1.29	1.41	1.24	1.48	1.17	1.55	1.12	1.63	1.06	1.70
35	1.30	1.42	1.25	1.48	1.19	1.55	1.13	1.63	1.07	1.70
36	1.31	1.43	1.26	1.49	1.20	1.56	1.15	1.63	1.09	1.70
37	1.32	1.43	1.27	1.49	1.21	1.56	1.16	1.62	1.10	1.70
38	1.33	1.44	1.28	1.50	1.23	1.56	1.17	1.62	1.12	1.70
39	1.34	1.44	1.29	1.50	1.24	1.56	1.19	1.63	1.13	1.69
40	1.35	1.45	1.30	1.51	1.25	1.57	1.20	1.63	1.15	1.69
45	1.39	1.48	1.34	1.53	1.30	1.58	1.25	1.63	1.21	1.69
50	1.42	1.50	1.38	1.54	1.34	1.59	1.30	1.64	1.26	1.69
55	1.45	1.52	1.41	1.56	1.37	1.60	1.33	1.64	1.30	1.69
60	1.47	1.54	1.44	1.57	1.40	1.61	1.37	1.65	1.33	1.69
65	1.49	1.55	1.46	1.59	1.43	1.62	1.40	1.66	1.36	1.69
70	1.51	1.57	1.48	1.60	1.45	1.63	1.42	1.66	1.39	1.70
75	1.53	1.58	1.50	1.61	1.47	1.64	1.45	1.67	1.42	1.70
80	1.54	1.59	1.52	1.62	1.49	1.65	1.47	1.67	1.44	1.70
85	1.56	1.60	1.53	1.63	1.51	1.65	1.49	1.68	1.46	1.71
90	1.57	1.61	1.55	1.64	1.53	1.66	1.50	1.69	1.48	1.71
95	1.58	1.62	1.56	1.65	1.54	1.67	1.52	1.69	1.50	1.71
100	1.59	1.63	1.57	1.65	1.55	1.67	1.53	1.70	1.51	1.72

SOURCE: J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression," *Biometrika*, vol. 38 (1951), pp. 159-177. Reprinted with permission of the authors and the Trustees of Biometrika.

إجابة الأسئلة

الفصل الأول

$$\bar{X}=3, \sum_{i=1}^n (X_i - \bar{X}) = -3 + 2 + 3 - 2 = 0. \quad -١$$

$$\sum_{i=1}^2 (aX_i + bY_i + cZ_i) = (aX_1 + bY_1 + cZ_1) + \dots + (aX_n + bY_n + cZ_n) \quad -٢$$

$$= (aX_1 + \dots + aX_n) + (bY_1 + \dots + bY_n) + (cZ_1 + \dots + cZ_n)$$

$$= \sum_{i=1}^n aX_i + \sum_{i=1}^n bY_i + \sum_{i=1}^n cZ_i$$

$$= a \sum_{i=1}^n X_i + b \sum_{i=1}^n Y_i + c \sum_{i=1}^n Z_i$$

$$\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}) = \sum_{i=1}^n [X_i(Y_i - \bar{Y}) - \bar{X}(Y_i - \bar{Y})] \quad -٣$$

$$= \sum_{i=1}^n X_i(Y_i - \bar{Y}) - \sum_{i=1}^n \bar{X}(Y_i - \bar{Y}).$$

ولكن، بما أن \bar{X} ثابت، فإنه يمكننا كتابة الحد الأخير على النحو

$$\bar{X} \sum_{i=1}^n (Y_i - \bar{Y}) = 0$$

الفصل الثاني

$$1. \hat{a} = \bar{Y} - \hat{b}\bar{X} = \sum_{i=1}^n \frac{Y_i}{n} - \bar{X} \sum_{i=1}^n \frac{(X_i - \bar{X})Y_i}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad -1$$

$$\text{حيث: } A = \sum_{i=1}^n (X_i - \bar{X})^2$$

$$d = \sum_{i=1}^n \frac{Y_i}{n} - \frac{\bar{X}}{A} \sum_{i=1}^n (X_i - \bar{X})Y_i = \sum_{i=1}^n \left[\frac{Y_i}{n} - \frac{\bar{X}}{A} (X_i - \bar{X}) \right] Y_i.$$

وبما أن $W_i = (X_i - \bar{X}) / A$ عندئذ تتحقق الإجابة المطلوبة.

٢- (أ) المعادلات الطبيعية هي :

$$\sum Y_i = n\hat{a}_0 + \hat{b} \sum X_i,$$

$$\sum X_i Y_i = \hat{a}_0 \sum X_i + \hat{b} \sum X_i^2.$$

وتعطينا حسابات القيم المشاهدة لـ X_i و Y_i النتائج.

$$\sum_{i=1}^5 Y_i = 30, \sum_{i=1}^5 X_i = 12, \sum_{i=1}^5 X_i^2 = 34, \sum_{i=1}^5 X_i Y_i = 74, N = 5$$

(ب) تعطينا المعادلات الطبيعية :

$$30 = 5\hat{a} + 12\hat{b},$$

$$74 = 12\hat{a} + 34\hat{b}.$$

وبحل هذه المجموعة من المعادلات، نحصل على :

$$\hat{a} = 5.076, \quad \hat{b} = 0.385.$$

$$\hat{\sigma}^2 = \frac{\sum (Y_i - \hat{Y}_i)^2}{n-2} = 3.077.$$

٣- هذه العبارة خاطئة لأنها تهمل وجود الخطأ العشوائي، u_i ، الذي يأخذ قيماً موجبة وقيماً سالبة. ولكن قيمته المتوقعة $E(u_i)$ هي الصفر. العلاقة $C_i = a + bY_i$ ليست علاقة مؤكدة، ولكنها، في الأصح، علاقة وسط.

٤- نشق أولاً μ_Y :

$$E(Y) = E(5 - 3X) = 5 - 3E(X) = 5 - 3\mu_X = \mu_Y.$$

ولذا، يكون التباين :

$$\begin{aligned}\sigma_{X,Y} &= E(Y - \mu_Y)(X - \mu_X) = E(5 - 3X - 5 + 3\mu_X)(X - \mu_X) \\ &= E[-3(X - \mu_X)^2] = -3\sigma_X^2.\end{aligned}$$

ويكون تباين Y هو $E(Y - \mu_Y)^2 = 9\sigma_X^2$ ، وهكذا، فإن $\sigma_Y = 3\sigma_X$. لذا، يكون معامل الارتباط هو :

$$\rho_{XY} = \frac{-3\sigma_X^2}{3\sigma_X\sigma_X} = -1$$

٥- نستخدم هنا العلاقة الأساسية لتباين مجموع المتغيرات العشوائية. توضح هذه العلاقة أنه إذا كان $Y = a_0X + a_1X_1 + \dots + a_nX_n$ وكانت (X_n, \dots, X_1) مستقلة، حيث :

$$\text{var}(Y) = a_1^2\sigma_{X_1}^2 + a_2^2\sigma_{X_2}^2 + \dots + a_n^2\sigma_{X_n}^2$$

وبتطبيق هذه العلاقة للمسألة نحصل على :

$$\text{var}(Y) = 4 + 27 + 500 = 531$$

٦- (أ) يمكن صياغة التحليل على النحو :

$$Y_i = a + b(T_{ci} - T_{si}) + u_i,$$

حيث إن :

$$Y_i = \text{متوسط دخل الاسرة في المدينة } i$$

$$T_{ci} = \text{معدل الضريبة في المدينة } i$$

$$T_{si} = \text{معدل الضريبة في ضواحي المدينة } i, \text{ وأخيراً}$$

$$u_i = \text{الخطأ العشوائي.}$$

(ب) تكوين مختلف للافتراض نفسه هو :

$$Y_i = a + b \frac{T_{ci}}{T_{si}} + u_i$$

في كلتا الحالتين، تكون $b < 0$ ، أي أنه إذا كان T_c مرتفعاً بالنسبة لـ T_s ، نتوقع أن تقيم الأسرة متوسطة الدخل في الضواحي ومرتفعته، وهكذا فإن متوسط الدخل العائلي لتلك الأسرة التي تبقى في المدينة، Y_i ، سيكون منخفضاً.

٧- بالتعويض من (2) في (1)، نحصل على :

$$Y_i = (a_1 + a_2) + (b_1 + b_2)X_i + \varepsilon_i.$$

وهكذا، فإن تأثير X_i على القيمة المتوسطة لـ Y_i سيكون $(b_1 + b_2)$. تبين هذه المشكلة أنه يمكننا اعتبار نموذجنا للانحدار الثنائي المعتاد مشتقاً من نموذج آخر يكون فيه الخطأ العشوائي مرتبطاً خطياً بالمتغير المستقل كما في (2).

٨- نعم، ولييان ذلك، عوض عن الخطأ العشوائي في المعادلة (1) من السؤال السابق للحصول على :

$$Y_i = (a_1 + a_2) + b_1X_i + (b_2X_i^2 + \varepsilon_i).$$

وهكذا، فإن النموذج الذي يربط Y_i و X_i سيحتوي على $w_i = b_2X_i^2 + \varepsilon_i$ باعتباره عشوائياً. ومن الواضح أن w_i لن تكون لها قيمة متوسطة صفرية. أكثر من ذلك، فإنه طالما أن X_i و X_i^2 مرتبطان ارتباطاً واضحاً. فإن w_i ستكون مرتبطة مع المتغير المستقل X_i أيضاً.

وإذا كانت قيم X_i مرتبطة عند نقاط زمنية مختلفة، لذلك، ستكون w_i كذلك، ولذلك فإن $\text{cov}(w_i, w_s) \neq 0$. أخيراً، طالما أن قيم w_i تعتمد جزئياً على قيم X_i ، فإن تباین w_i لن يكون متساوياً في كل فترة زمنية.

٩- دع A_t عمر الطفل في الزمن t و H_t طوله مقاساً بالبوصات. حيثُذ، يمكننا أن نفترض :

$$H_t = a + bA_t + u_t,$$

حيث نتوقع $b > 0$. وأحد أوجه القصور لهذه العلاقة هو أنها تكون صحيحة، فقط، لعدد محدود من السنوات، أي أنه، على الرغم من أن الطفل سوف يتقدم في العمر سنة بعد أخرى، فإن ارتفاعه لن يتجاوز حداً معيناً أعلى.

١- (أ) سيأخذ نموذج الانحدار الشكل :

$$Y_t = a + bX_t^m + (u_t - b\varepsilon_t).$$

(ب) نعم، سيكون الخطأ العشوائي مرتبطاً بالمتغير المستقل X_t^m . ولتوضيح ذلك، لاحظ أن :

$$\begin{aligned} E[X_t^m(u_t - b\varepsilon_t)] &= E(X_t^m u_t) - bE(X_t^m \varepsilon_t) \\ &= 0 - b\sigma_\varepsilon^2 \neq 0, \end{aligned}$$

طالما أن $(X_t^m \varepsilon_t) = X_t \varepsilon_t + \varepsilon_t^2$ وأن X_t و ε_t مستقلان .

الفصل الثالث

١- دع S متوسط مقياس الذكاء I-Q لجميع الافراد في المجتمع، حينئذ، نختبر $H_0 : S = 100$ مقابل $H_1 : S > 100$. وطالما أننا مهتمون، فقط، بالحالة $S \geq 100$ فيمكننا استخدام اختبار الذيل الواحد لـ t مع درجات حرية قدرها $(100-1=99)$. نجد أن القيمة الدنيا لـ S هي $[\hat{S} - (t_{n-1; 0.95}) \hat{\sigma}_s = 110 - 1.65(2)]$ وطالما أن الحد الأدنى لـ S أعلى من 100، فإننا نرفض $H_0 : S = 100$ ونقبل $H_1 : S > 100$.

٢- دع \bar{X}_{20} و \bar{X}_{30} متوسطات العينة والمبنية على حجم 20 و 30 على الترتيب. حينئذ، إذا كانت العيتان كلتاهما قد سحبتا من المجتمع نفسه، حيث $\mu = E(\bar{X}_{20}) = E(\bar{X}_{30})$ هي القيمة المتوسطة للمجتمع. وهكذا يكون كلا المقدرين غير متحيزين. ولكننا سنفضل \bar{X}_{30} لأن تباينها سيكون أصغر. لذلك فإن استخدامها سيؤدي إلى فترات ثقة أضيق واختبارات افضل للفرضيات.

٣- نعرف من النتائج أن $\hat{\alpha} = 0.125$. ولذا، قد نختبر الافتراض $H_0 : \alpha = 0$ مقابل $H_1 : \alpha \neq 0$ ، ويتم ذلك في معادلتنا عن طريق اختبار $H_0 : (1 - \alpha) = 1$ مقابل $H_1 : (1 - \alpha) \neq 1$ وتكون القيمة المطلقة لنسبة t هي :

$$\left| \frac{0.875 - 1.00}{0.15} \right| = \frac{0.125}{0.15} = 0.83,$$

التي تكون أقل كثيرا من 2. كذلك، نقبل الفرض العدمي ونستنتج أن التغير في الساعات لا يتناسب مع الانحراف عن 40.

٤- دع h القيمة المتوقعة للارتفاع. نختبر بعد ذلك $H_0 : h = 70$ مقابل $H_1 : h < 70$ وطالما أننا مهتمون، فقط، بـ $h \leq 70$ ، فيمكننا أن نستخدم اختبار الذيل الواحد. وهنا نجد أن الحد الأعلى لقيمة h هي $(\hat{h} + 1.65\sigma_h = 68 + 1.65(2) = 71.3)$ وطالما أن $70 \leq 71.3$ فإننا نقبل $H_0 : h = 70$.

٥- يسهل لنا افتراض الطبيعية بناء فترات الثقة واختبار الفرضيات، وقد وضح ذلك في الكتاب، غير أن افتراض الطبيعية ليس شرطا ضروريا لبناء فترات الثقة، أو اختبار الفرضيات، ولكن، بدون افتراض الطبيعية، ستصبح هذه المهام أكثر صعوبة.

٦- ستكون فرضية العدم : إن القيمة المتوسطة لأطول الناس في الغرب 67 بوصة، وتكون الفرضية البديلة أن القيمة المتوسطة لأطوالهم يزيد على 67 بوصة. ونقع في النوع الأول من الخطأ عندما ندفع للاعتقاد بأنهم أطول وهم، في الحقيقة، غير ذلك. ويتبع عن ذلك أن إعادة تجهيز غير ضرورية للتجربة ستحدث. إضافة إلى ذلك، فإن معاطف من الحجم غير الصحيح ستتج. ونقع في الخطأ من النوع الثاني عندما ندفع للاعتقاد بأنهم ليسوا أطولا وهم، في الحقيقة، كذلك. ويترتب على ذلك أنه سيتم إنتاج معاطف من الحجم غير المناسب. ويتضمن هذا أن النوع الأول من الخطأ قد يكون، في حالتنا هذه، أكثر تكلفة من النوع الثاني من الخطأ.

٧- ليس النموذج الخطي مقيلاً بتلك الدرجة كما يظهر لك، لأن الاستخدام الجيد للتحويلات المختلفة يمكننا من تحويل تشكيلة عريضة من العلاقات غير الخطية إلى أشكال خطية. وبجعل $Z_t = 1/(1 - X_t)$ ، تكون مصفوفة المشاهدات

هي:

Y_t	X_t	Z_t
1	0	1
10	0.1	1.11
12	0.5	2

٨- (أ) نختبر الفرضية (عند مستوى معنوية 5%) عن طريق ملاحظة هل تزداد القيمة المطلقة لنسبة t عن 2 بصورة معنوية، وطالما أنها تحقق ذلك فإننا نرفض فرضية العدم.

(ب) الانحرافات المعيارية هي :

$$\hat{\sigma}_a = \frac{15}{3.1} \doteq 4.84, \quad \hat{\sigma}_b = \frac{0.81}{18.7} \doteq 0.043,$$

(ج) وتكون 95% فترة ثقة هي :

$$0.81 \pm (t_{n-2, 0.975}) 0.043 = 0.81 \pm 0.091.$$

٩- قد يمكن التعبير عن فرضية أن الطلب على الرعاية الاجتماعية، D_t ، يرتبط بمعدل الإعانة B_t على النحو :

$$D_t = a_1 + a_2 B_t + u_{1t},$$

حيث نتوقع أن $a_2 > 0$. ويمكن التعبير عن فرضية أن معدل الإعانة يرتبط بالطلب على الرعاية الاجتماعية، من خلال الضغط السياسي، على النحو:

$$B_t = b_1 + b_2 D_{t-1} + u_{2t},$$

وذلك إذا افترضنا وجود فترة إبطاء بسبب العملية السياسية.

١٠- دع N_t عدد المنشآت التي تتوطن ولاية معينة. حيثئذ، فإن نموذج التوطن قد يمكن التعبير عنه على النحو :

$$N_t = a_1 + b_1 \left(\frac{T_{1t}}{T_{2t}} \right) + u_{1t},$$

حيث T_{1t} هي الضريبة في ولاية معينة، T_{2t} هي معدل الضريبة المتوسطة في الولايات المجاورة ولذلك، نتوقع أن يكون $b_1 < 0$. وبالمثل إذا جعلنا P_t يرمز إلى أحد مقاييس التلوث، حيثئذ فإن علاقة التلوث قد يمكن التعبير عنها على النحو :

$$P_t = a_2 + b_2 N_t + u_{2t},$$

ونتوقع أن يكون $b_2 > 0$.

١١- إذا قمنا بالحل للحصول على X_t ، يكون لدينا :

$$X_t = \frac{5 - Z_t}{3}.$$

وبالتعويض في نموذج الانحدار، يكون لدينا :

$$\begin{aligned} Y_t &= a + b \left(\frac{5 - Z_t}{3} \right) + u_t \\ &= \left(a + \frac{5}{3}b \right) - \frac{b}{3} Z_t + u_t. \end{aligned}$$

وهكذا يكون الحد الثابت هو $(a + 5/3 b)$ والميل هو $(-b/3)$.

الفصل الرابع

١- افتراضات النموذج تكون على النحو التالي :

- (أ) ١- تكون القيمة المتوقعة للخطأ العشوائي مساوية للصفر $E(u_t) = 0$.
- ٢- يكون تباين الخطأ العشوائي ثابتاً $E(u_t^2) = \sigma_u^2$.
- ٣- تكون قيمة الخطأ العشوائي لإحدى المشاهدات مستقلة عن قيمته لأي مشاهدة أخرى، لذلك يكون التغاير بين أي خطئين عشوائيين لأي مشاهدين u_t و u_s مساوياً للصفر $\text{cov}(u_t, u_s) = 0$.

٤ - الخطأ العشوائي مستقل عن كل واحد من المتغيرات المستقلة،

والقيم المبطة لها كافة، لذلك فإن $\text{cov}(u_t, X_{it}) = 0$.

٥ - لا يكون أي من المتغيرات المستقلة مولفاً خطياً من الآخرين.

(ب) المعادلات الطبيعية هي :

$$1. \sum Y_t = n\hat{a}_0 + \hat{a}_1 \sum X_{1t} + \hat{a}_2 \sum X_{2t},$$

$$2. \sum X_{1t}Y_t = \hat{a}_0 \sum X_{1t} + \hat{a}_1 \sum X_{1t}^2 + \hat{a}_2 \sum X_{1t}X_{2t},$$

$$3. \sum X_{2t}Y_t = \hat{a}_0 \sum X_{2t} + \hat{a}_1 \sum X_{2t}X_{1t} + \hat{a}_2 \sum X_{2t}^2.$$

نشتق المعادلة الطبيعية الأولى عن طريق جعل $\sum \hat{u}_t = 0$ وينظر هذا الافتراض

$E(u_t) = 0$. بينما تشتق المعادلة الطبيعية الثانية عن طريق جعل $\sum (X_{1t}\hat{u}_t) = 0$ ،

وينظر هذا الافتراض $E(X_{1t} u_t) = \text{cov}(X_{1t}, u_t) = 0$. وتشتق المعادلة الطبيعية

الثالثة عن طريق $\sum (X_{2t}\hat{u}_t) = 0$ ، وينظر الافتراض

$$E(X_{2t} u_t) = \text{cov}(X_{2t}, u_t) = 0$$

(ج)

$$10 = 100\hat{a}_0$$

$$30 = 35\hat{a}_1$$

$$20 = 3\hat{a}_2$$

$$\hat{a}_0 = \frac{1}{10}$$

$$\hat{a}_1 = \frac{6}{7}$$

$$\hat{a}_2 = \frac{20}{3}$$

٢- المعلومات التي لا يمكن تقديرها هي a_1 ، a_2 و a_3 . ويرجع ذلك إلى أن المتغير

المستقل الثالث $(X_{1t} - X_{2t})$ هو توليفة خطية من X_{1t} و X_{2t} . لذلك، يوجد لدينا

تعدد علاقات خطية تام. ويعني ذلك أن المعادلات الطبيعية المناظرة هي توليفة خطية من المعادلات الطبيعية المناظرة لـ X_{1t} و X_{2t} ، وهكذا، لا يمكننا، عموماً، أن نحل هذه المعادلات للحصول على مقدرات المعلمات. لاحظ أن $X_{2t}X_{1t}$ توليفة غير خطية، ومن ثم، لا يمثل مشكلة لنا. وبإعادة كتابة المعادلة على النحو :

$$Y_t = a_0 + (a_1 + a_3)X_{1t} + (a_2 - a_3)X_{2t} + a_4X_{1t}X_{2t} + \varepsilon_t,$$

نجد أنه يمكننا الحل للحصول على \hat{a}_0 ، $(\hat{a}_1 + \hat{a}_3)$ ، $(\hat{a}_2 - \hat{a}_3)$ و \hat{a}_4 .
٣- المعادلات الطبيعية هي :

$$\begin{aligned}\sum Y_t &= n\hat{b}_0 + \hat{b}_1 \sum X_{1t} + \hat{b}_2 \sum X_{2t}, \\ \sum X_{1t}Y_t &= \hat{b}_0 \sum X_{1t} + \hat{b}_1 \sum X_{1t}^2 + \hat{b}_2 \sum X_{1t}X_{2t}, \\ \sum X_{2t}Y_t &= \hat{b}_0 \sum X_{2t} + \hat{b}_1 \sum X_{2t}X_{1t} + \hat{b}_2 \sum X_{2t}^2.\end{aligned}$$

وتعطينا الحسابات، باستخدام القيم المشاهدة لكل من X_t و Y_t مايلي :

$$\sum X_{1t}X_{2t} = 12, \sum Y_t = 30, \sum X_{2t} = 9, \sum X_{1t} = 8, n = 5$$

$$\sum X_{1t}^2 = 16, \sum X_{2t}^2 = 19, \sum X_{1t}Y_t = 50, \sum X_{2t}Y_t = 51$$

وبإدخال هذه القيم المحسوبة في معادلتنا الطبيعية نحصل على :

$$30 = 5\hat{b}_0 + 8\hat{b}_1 + 9\hat{b}_2,$$

$$50 = 8\hat{b}_0 + 18\hat{b}_1 + 12\hat{b}_2,$$

$$51 = 9\hat{b}_0 + 12\hat{b}_1 + 19\hat{b}_2.$$

٤- يمكن التعبير عن نموذج الانحدار على النحو :

$$Y_i = a + b_1(T_{ci} - T_{si}) + b_2(C_{ci} - C_{si}) + b_3(H_{ci} - H_{si}) + b_4D_{ci} + u_i,$$

$$Y_i = \text{الدخل المتوسط في المدينة } i,$$

$$T_{ci} = \text{معدل الضريبة في المدينة } i,$$

$$T_{si} = \text{معدل الضريبة في الضاحية المناظرة } i,$$

$$C_{ci} = \text{معدل الجريمة في المدينة } i,$$

$$\begin{aligned}
C_{si} &= \text{معدل الجريمة في ضاحية المدينة } i, \\
H_{ci} &= \text{تكاليف السكن في المدينة } i, \\
H_{si} &= \text{تكاليف السكن في الضاحية المناظرة } i, \\
D_{ci} &= \text{الكثافة السكانية في المدينة } i \text{ (سكان/ميل مربع)}, \text{ وأخيراً} \\
u_i &= \text{الخطأ العشوائي.}
\end{aligned}$$

تعتمد أهمية البيانات المقطعية وبيانات السلاسل الزمنية على طبيعة المشكلة. وفي انحدارنا هذا، تكون البيانات المقطعية، على الأرجح، أكثر فائدة لأن معدلات الضرائب ومعدلات الجرائم، وتكاليف السكن والكثافة تميل للتغير تغيراً كبيراً عبر المدن أكثر منه عبر الزمن داخل مدينة معينة. وفي هذا النموذج، نتوقع أن تأخذ معلمات النموذج كافة قيماً سالبة.

٥- (أ) العلاقة الخطية التامة :

$$\bar{P}_i = \frac{\sum_{t=1}^k P_{it}}{k}$$

تتضمن أن المعادلة الطبيعية (k+1) تكون توليفة خطية من المعادلات الطبيعية الـ k الأولى. وبينما يكون لدينا معلمات عددها (k+3) لكي نقدرها، فإنه يتوافر لدينا، فقط، (k+2) من المعادلات الطبيعية المستقلة. وهكذا لا يمكننا، عموماً، أن نحل هذا النموذج من أجل الحصول على قيم فريدة لمقدرات المعلمات.

(ب) بالتعويض عن $P_t = \sum_{i=1}^k (p_{it}) / k$ في معادلة الطلب وتجميع الحدود نحصل على :

$$\begin{aligned}
D_{it} &= a_0 + P_{1t} \left(a_1 + \frac{b}{k} \right) + P_{2t} \left(a_2 + \frac{b}{k} \right) + \dots \\
&\quad + P_{kt} \left(a_k + \frac{b}{k} \right) + cY_t + u_{it}.
\end{aligned}$$

وهكذا سنكون قادرين على تقدير a_0 و $(a_1 + b/4)$ و c حيث إن $i=1, \dots, k$.

٦- (أ) لا تعاني هذه المعادلة من تعدد علاقات خطية تام طالما أن X_t و X_t^2 ليسا مرتبطين ارتباطا تاما، وتكون المعادلات الطبيعية:

$$\begin{aligned}\sum Y_t &= n\hat{b}_0 + \hat{b}_1 \sum X_t + \hat{b}_2 \sum X_t^2, \\ \sum Y_t X_t &= \hat{b}_0 \sum X_t + \hat{b}_1 \sum X_t^2 + \hat{b}_2 \sum X_t^3, \\ \sum Y_t X_t^2 &= \hat{b}_0 \sum X_t^2 + \hat{b}_1 \sum X_t^3 + \hat{b}_2 \sum X_t^4.\end{aligned}$$

نلاحظ أن هذه المعادلات الثلاثة مستقلة خطيا، ولذا، يمكننا أن نحلها للحصول على \hat{b}_0 ، \hat{b}_1 و \hat{b}_2 .

(ب)

$$\begin{aligned}4 &= 4\hat{b}_0 + 8\hat{b}_1 + 30\hat{b}_2, \\ 23 &= 8\hat{b}_0 + 30\hat{b}_1 + 134\hat{b}_2, \\ 107 &= 30\hat{b}_0 + 134\hat{b}_1 + 642\hat{b}_2.\end{aligned}$$

الفصل الخامس

١- باستخدام التحويل اللوغاريتمي للدالة، نحصل على :

$$Q'_t = B + aL'_t + bK'_t + u_t,$$

حيث :

$$Q'_t = \text{لوغاريتم } Q,$$

$$B = \text{لوغاريتم } (1/A),$$

$$L'_t = \text{لوغاريتم } L, \text{ وأخيرا}$$

$$K'_t = \text{لوغاريتم } K.$$

نقدر B ، a و b ثم نأخذ $\hat{A} = e^{-\hat{B}}$. نلاحظ أن \hat{A} سيكون متحيزا ولكن متسقا

٢- نموذج الانحدار هو :

$$I_t = a_0 + b_1 r_t + b_2 D_t + u_t,$$

حيث :

إذا كان الرئيس من الديمقراطيين في الزمن t ، $0 = D_t$

إذا كان الرئيس من الجمهوريين ، $1 = D_t$

معدل الفائدة ، r_t

الخطأ العشوائي . u_t

٣- النموذج هو :

$$C_t = a_0 + a_1 Z_{1t} + a_2 Z_{2t} + a_3 Z_{3t} + u_t,$$

$$Z_{1t} = F_t Y_t,$$

$$Z_{2t} = Y_t^{1/2},$$

$$Z_{3t} = \frac{1}{A_t}.$$

٤- لنأخذ الانتقال المحتمل للدالة في الحسبان، ندخل متغيراً صورياً في النموذج،

ومن ثم، يصبح نموذج الانحدار :

$$C_t = a + b Y_t + c D_t + u_t,$$

حيث

$0 = D_t$ إذا كان المستهلك i يعيش في منطقة حضرية،

$1 = D_t$ إذا كان يعيش في منطقة أخرى.

وهكذا، إذا كان المستهلك i يقيم بمنطقة ريفية يكون الانحدار

$C_t = (a+c) + b Y_t + u_t$ بينما إذا كان يقيم في منطقة حضرية فإن الدالة تصبح

$$C_t = a + b Y_t + u_t.$$

٥- (أ) نموذج الانحدار هو :

$$I_t = b_0 + b_1 r_t + b_2 \Pi_t + b_3 \Delta S_t + u_t,$$

حيث

I_t = الإنفاقات الاستثمارية،

Π_t = معدل الربح،

ΔS_t = التغير في المبيعات،

$$r_t = \text{معدل الفائدة.}$$

$$u_t = \text{الخطأ العشوائي}$$

(ب) المشكلة في تقدير الانحدار الحالي هي وجود تعدد العلاقات الخطية. وبالتحديد إذا كان معدل الأرباح هو ١٥٪ في كل فترة زمنية، فلن نكون قادرين على تقدير b_0 و b_2 .

٦- (أ) يكون الشكل غير المقيد من النموذج هو :

$$I_t = a_0 + a_1 r_t + b_0 S_t + b_1 S_{t-1} + \dots + b_7 S_{t-7} + u_t.$$

(ب) شكل آلمون للنموذج، باستخدام $b_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2$ ، هو :

$$I_t = a_0 + a_1 r_t + \alpha_0 Z_{1t} + \alpha_1 Z_{2t} + \alpha_2 Z_{3t} + u_t.$$

حيث :

$$Z_{1t} = \sum_{i=0}^7 s_{t-i}, \quad Z_{2t} = \sum_{i=1}^7 i s_{t-i}, \quad Z_{3t} = \sum_{i=1}^7 i^2 s_{t-i}$$

(ج) المعادلات الطبيعية هي :

$$\begin{aligned} \sum_{t=0}^n I_t &= n\hat{a}_0 + \hat{a}_1 \sum_{t=0}^n r_t + \hat{\alpha}_0 \sum_{t=0}^n Z_{1t} \\ &+ \hat{\alpha}_1 \sum_{t=0}^n Z_{2t} + \hat{\alpha}_2 \sum_{t=0}^n Z_{3t}, \\ \sum_{t=0}^n r_t I_t &= \hat{a}_0 \sum_{t=0}^n r_t + \hat{a}_1 \sum_{t=0}^n r_t^2 + \hat{\alpha}_0 \sum_{t=0}^n r_t Z_{1t} \\ &+ \hat{\alpha}_1 \sum_{t=0}^n r_t Z_{2t} + \hat{\alpha}_2 \sum_{t=0}^n r_t Z_{3t}, \\ \sum_{t=0}^n Z_{1t} I_t &= \hat{a}_0 \sum_{t=0}^n Z_{1t} + \hat{a}_1 \sum_{t=0}^n Z_{1t} r_t + \hat{\alpha}_0 \sum_{t=0}^n Z_{1t}^2 \\ &+ \hat{\alpha}_1 \sum_{t=0}^n Z_{1t} Z_{2t} + \hat{\alpha}_2 \sum_{t=0}^n Z_{1t} Z_{3t}, \\ \sum_{t=0}^n Z_{2t} I_t &= \hat{a}_0 \sum_{t=0}^n Z_{2t} + \hat{a}_1 \sum_{t=0}^n Z_{2t} r_t + \hat{\alpha}_0 \sum_{t=0}^n Z_{2t} Z_{1t} \\ &+ \hat{\alpha}_1 \sum_{t=0}^n Z_{2t}^2 + \hat{\alpha}_2 \sum_{t=0}^n Z_{2t} Z_{3t}, \end{aligned}$$

$$\begin{aligned} \sum_{t=0}^m Z_{3t} I_t = \hat{\alpha}_0 \sum_{t=0}^m Z_{3t} + \hat{\alpha}_1 \sum_{t=0}^m Z_{3t} I_t + \hat{\alpha}_0 \sum_{t=0}^m Z_{3t} Z_{1t} \\ + \hat{\alpha}_1 \sum_{t=0}^m Z_{3t} Z_{2t} + \hat{\alpha}_2 \sum_{t=0}^m Z_{3t}^2. \end{aligned}$$

٧- (أ) سيكون تقدير b_2 على النحو :

$$\begin{aligned} \hat{b}_2 &= \hat{\alpha}_0 + 2\hat{\alpha}_1 + 4\hat{\alpha}_2 + 8\hat{\alpha}_3 + 16\hat{\alpha}_4 \\ &= 1 + 6 + 20 + 32 - 160 = -101. \end{aligned}$$

(ب) وبالتعويض عن الـ b 's في النموذج الأصلي، نحصل على :

$$Y_t = a + \alpha_0 Z_{1t} + \alpha_1 Z_{2t} + \alpha_2 Z_{3t} + \alpha_3 Z_{4t} + \alpha_4 Z_{5t} + u_t,$$

حيث :

$$\begin{aligned} Z_{1t} &= \sum_{i=0}^6 X_{t-i}, & Z_{2t} &= \sum_{i=0}^6 i X_{t-i}, \\ Z_{3t} &= \sum_{i=1}^6 i^2 X_{t-i}, & Z_{4t} &= \sum_{i=0}^6 i^3 X_{t-i}, \\ Z_{5t} &= \sum i^4 X_{t-i}. \end{aligned}$$

وحيث إن :

$$\sum_{i=0}^6 b_i = 7\alpha_0 + \sum_{i=0}^6 i\alpha_1 + \sum_{i=0}^6 i^2\alpha_2 + \sum_{i=0}^6 i^3\alpha_3 + \sum_{i=0}^6 i^4\alpha_4 = 1,$$

فإنه يمكننا أن نوجد α_0 على النحو :

$$\alpha_0 = \frac{(1 - 21\alpha_1 - 91\alpha_2 - 441\alpha_3 - 2275\alpha_4)}{7}.$$

وبالتعويض عن α_0 في الانحدار السابق، نحصل على :

$$Y_t = a + \alpha_1 Q_{1t} + \alpha_2 Q_{2t} + \alpha_3 Q_{3t} + \alpha_4 Q_{4t} + u_t,$$

حيث :

$$Y_t^* = \left(Y_t - \frac{Z_{1t}}{7} \right),$$

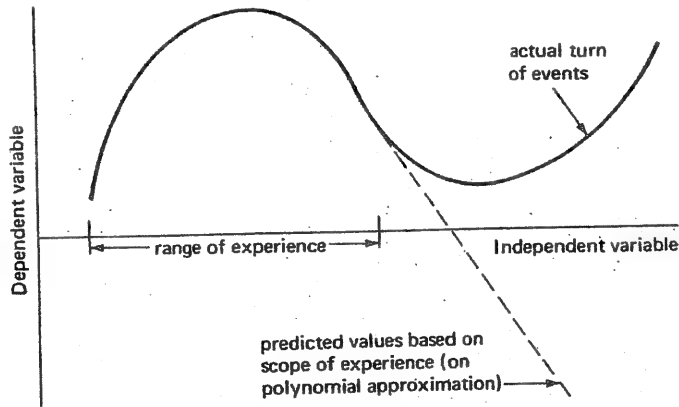
$$Q_{1t} = Z_{2t} - \frac{21Z_{1t}}{7},$$

$$Q_{2t} = Z_{3t} - \frac{91Z_{1t}}{7},$$

$$Q_{3t} = Z_{4t} - \frac{441Z_{1t}}{7},$$

$$Q_{4t} = Z_{5t} - \frac{2275Z_{1t}}{7}$$

٨- معظم الدوال التي يعالجها الاقتصاديون يمكن تقريبها بمتعدد الحدود. وتحدد درجة متعدد الحدود بنطاق الخبرة أو بعدد المتغيرات المتضمنة في الدالة. بالنسبة للحالات الجديدة، خارج نطاق الخبرة، فإن استخدام متعدد الحدود بدرجة مختلفة قد يكون ملائماً. لذلك، فقد لا يكون من الملائم أن نستخدم معادلتنا المقدرة لأغراض التنبؤ هذه. يتضح هذا من الشكل التالي :



٩- بتحويل المعادلة إلى معادلة مبطئة والضرب في λ ثم طرحها من المعادلة الأصلية يكون لدينا :

$$Y_t = (a_0 - \lambda a_0) + a_1 X_t - a_1 \lambda X_{t-1} + \lambda Y_{t-1} + b_0 Z_t + u_t - \lambda u_{t-1}.$$

١٠- إذا كانت $b_5 = 3$ ، يكون لدينا $\alpha_0 + 5\alpha_1 + 25\alpha_2$. لذلك، قد يمكننا حل هذه المعادلة للحصول على $\alpha_0 = 3 - 5\alpha_1 - 25\alpha_2$. ويكون الشكل غير المقيد لنموذج آلون هو:

$$Y_t = b + \alpha_0 Z_{0t} + \alpha_1 Z_{1t} + \alpha_2 Z_{2t} + u_t, \quad \text{حيث}$$

$$Z_{0t} = \sum_{i=0}^{10} X_{t-i}, \quad Z_{1t} = \sum_{i=1}^{10} i X_{t-i}, \quad Z_{2t} = \sum_{i=1}^{10} i^2 X_{t-i}.$$

وبالتعويض عن α_0 ، نحصل على الشكل المقيد :

$$Y_t^* = b + \alpha_1 Q_{1t} + \alpha_2 Q_{2t} + u_t, \quad \text{حيث}$$

$$Y_t^* = Y_t - 3Z_{0t}, \quad Q_{1t} = (Z_{1t} - 5Z_{0t}), \quad \text{وأيضا،}$$

$$Q_{2t} = (Z_{2t} - 25Z_{0t}).$$

١١- دع

$$D_t = \begin{cases} 1 & \text{في الحالات الأخرى} \\ 0 & \text{if } r_t > 0.05, \end{cases}$$

حيث، يكون نموذجنا للانحدار هو :

$$C_t = a + b_1 Y_t + b_2 (D_t r_t) + u_t.$$

١٢- دع :

$$\log Y_t = Y_t^*,$$

$$e^{X_{1t}} = Z_{1t},$$

$$\frac{1}{1 + X_{1t} X_{2t}} = Z_{2t}.$$

حيث، يمكن كتابة نموذج الانحدار على النحو :

$$Y_t^* = a_0 + a_1 Z_{1t} + a_2 Z_{2t} + u_t.$$

وتكون المعادلة الطبيعية هي :

$$\begin{aligned}\sum Y_i^* &= n\hat{a}_0 + \hat{a}_1 \sum Z_{1i} + \hat{a}_2 \sum Z_{2i}, \\ \sum Z_{1i} Y_i^* &= \hat{a}_0 \sum Z_{1i} + \hat{a}_1 \sum Z_{1i}^2 + \hat{a}_2 \sum Z_{1i} Z_{2i}, \\ \sum Z_{2i} Y_i^* &= \hat{a}_0 \sum Z_{2i} + \hat{a}_1 \sum Z_{1i} Z_{2i} + \hat{a}_2 \sum Z_{2i}^2.\end{aligned}$$

الفصل السادس

١- الخطوة الأولى: احسب \hat{a} و \hat{b}

$$\hat{b} = \frac{\sum_{i=1}^{15} (X_i - \bar{X}) Y_i}{\sum_{i=1}^{15} (X_i - \bar{X})^2} = \frac{255}{280} = 0.91.$$

$$\hat{a} = \bar{Y} - \hat{b}\bar{X} = -0.28.$$

ومن ثم فإن

$$\hat{Y}_i = -0.28 + 0.91 X_i, \quad \hat{u}_i = Y_i - (-0.28 + 0.91 X_i).$$

الخطوة الثانية: احسب \hat{u}_i^2 و $(\hat{u}_i - \hat{u}_{i-1})^2$:

Y_i	\hat{u}_i^2	$(\hat{u}_i - \hat{u}_{i-1})^2$
0.63	1.876	-
1.54	0.211	0.828
2.45	0.203	0.828
3.36	5.570	3.648
4.27	1.612	1.188
5.18	0.032	1.188
6.09	0.008	0.008
7.00	1.000	0.828
7.91	4.368	9.548
8.82	1.392	0.828
9.73	0.073	0.828
10.64	1.850	1.188
11.55	11.903	4.369
12.46	6.052	34.928
13.37	5.617	0.008
	41.767	60.213

$$\sum \hat{u}_t^2 = 41.767,$$

$$\sum (\hat{u}_t - \hat{u}_{t-1})^2 = 60.213.$$

ولذلك، تكون $d = 60.213 / 41.767 = 1.44$ ، وهي أكبر من الحد الأعلى 1.23 من الجدول الإحصائي رقم ٤. ولذا، نرفض وجود الارتباط الذاتي. ٢- (أ) من المنطقي أن نجادل بوجود اختلاف تباين، لأن من الصعب الاعتقاد أنه بينما ينمو الناتج على مدى الزمن، فإن تباين أحد مكوناته، الخطأ العشوائي، لا ينمو.

(ب) يؤدي حذف K_t من الانحدار إلى إيجاد مقدرات متحيزة لكل من a و b لأن الخطأ العشوائي في المعادلة الناتجة سيكون: $u_t^* = cK_t + u_t$. وستكون القيمة المتوسطة لهذا الخطأ العشوائي، عموماً، غير مساوية الصفر، وسيكون مرتبطاً بالمتغير المستقل L_t . نتوقع تحيزاً موجباً في \hat{b} لأن L_t ستضم أثرها إلى أثر رأس المال على الناتج معاً.

٣- (أ) بتجميع المعادلة (1) والقسمة على N ، نحصل على:

$$\sum_{i=1}^N \frac{C_{it}}{N} = a + b_1 \sum_{i=1}^N \frac{Y_{it}}{N} + b_2 \sum_{i=1}^N \frac{Y_{it}^2}{N} + \frac{\sum u_{it}}{N}.$$

وباستخدام تعريفاتنا، نحصل على:

$$C_t = a + b_1 Y_t + b_2 \sum_{i=1}^N \frac{Y_{it}^2}{N} + u_t \quad \text{والآن}$$

$$\begin{aligned} \sum_{i=1}^N Y_{it}^2 &= \sum_{i=1}^N (Y_{it} - Y_t + Y_t)^2 \\ &= \sum_{i=1}^N (Y_{it} - Y_t)^2 + \sum_{i=1}^N Y_t^2 + 2 \sum_{i=1}^N Y_t (Y_{it} - Y_t). \end{aligned}$$

ويساوي الحد الأخير :

$$2 \sum_{i=1}^N Y_t (Y_{it} - Y_t) = Y_t \sum_{i=1}^N (Y_{it} - Y_t) = 0$$

الصفر طالما أن Y_t هي القيمة المتوسطة لـ Y_{it} . لاحظ، أيضا، أن:

$$\sum_{i=1}^N Y_{it}^2 = N Y_t^2. \quad \text{حيث}$$

وهكذا، فإن نموذجنا الكلي للاقتصاد ينبغي أن يصبح بعد القسمة على N :

$$C_t = a + b_1 Y_t + b_2 Y_t^2 + b_3 s_t^2 + u_t,$$

حيث

$$s_t^2 = \sum_{i=1}^N \frac{(Y_{it} - Y_t)^2}{N}.$$

هذا الحد (s_t^2) هو مقياس للتغير في الدخل عبر المجتمع، لذلك، فإن النموذج الكلي، كما يظهر في المعادلة (3) يكون قد صيغ صياغة غير صحيحة. (ب) مصفوفة المشاهدات ستكتب بالطريقتين الأساسيتين التاليتين :

t	C_{it}	Y_{it}
1	C_{11}	Y_{11}
1	C_{21}	Y_{21}
1	C_{31}	Y_{31}
2	C_{12}	Y_{12}
2	C_{22}	Y_{22}
2	C_{32}	Y_{32}

t	C_{it}	Y_{it}
1	C_{11}	Y_{11}
2	C_{12}	Y_{12}
1	C_{21}	Y_{21}
2	C_{22}	Y_{22}
1	C_{31}	Y_{31}
2	C_{32}	Y_{32}

(ج) سيكون لدينا، عموما، تحيز التجميع Aggregation bias لأن متوسط الشكل غير الخطي للمتغير لن يعادل الشكل غير الخطي لمتوسط ذلك المتغير. على سبيل المثال، رأينا أن $\sum Y_{it}^2 / N \neq Y_t^2$ حيث Y_t هو متوسط Y_{it} . وبصورة أعم سيكون لدينا : $\sum_{i=1}^N f(X_{it}) / N \neq f(X_t)$ حيث X_t هي القيمة المتوسطة لـ X_{it} عبر المجتمع.

٤- (أ) المعادلات الطبيعية هي :

$$\begin{aligned}
(1) \quad \sum M_{dt} &= n\hat{b}_0 + \hat{b}_1 \sum i_t + \hat{b}_2 \sum i_{(t-1)} + \hat{b}_3 \sum \Delta i_t, \\
(2) \quad \sum i_t M_{dt} &= \hat{b}_0 \sum i_t + \hat{b}_1 \sum i_t^2 + \hat{b}_2 \sum i_t i_{(t-1)} + \hat{b}_3 \sum i_t \Delta i_t, \\
(3) \quad \sum i_{(t-1)} M_{dt} &= \hat{b}_0 \sum i_{t-1} + \hat{b}_1 \sum i_t i_{(t-1)} + \hat{b}_2 \sum i_{(t-1)}^2 + \hat{b}_3 \sum i_{t-1} \Delta i_t, \\
(4) \quad \sum \Delta i_t M_{dt} &= \hat{b}_0 \sum \Delta i_t + \hat{b}_1 \sum \Delta i_t i_t + \hat{b}_2 \sum \Delta i_t i_{t-1} + \hat{b}_3 \sum \Delta i_t^2
\end{aligned}$$

ونبتذكر أن $\Delta = i_t - i_{t-1}$ ، يمكننا إعادة كتابة المعادلة (4) على النحو :

$$\begin{aligned}
(5) \quad \sum (i_t - i_{t-1}) M_{dt} &= \hat{b}_0 \sum i_t - \hat{b}_0 \sum i_{t-1} + \hat{b}_1 \sum i_t^2 - \hat{b}_1 \sum i_t i_{t-1} \\
&\quad + \hat{b}_2 \sum i_t i_{t-1} - \hat{b}_2 \sum i_{t-1}^2 + \hat{b}_3 \sum i_t^2 \\
&\quad + \hat{b}_3 \sum i_{t-1}^2 - 2\hat{b}_3 \sum i_t i_{t-1}.
\end{aligned}$$

يتضح لنا من فحص بسيط أن المعادلة (5) تساوي المعادلة (2) مطروحاً منها المعادلة (3). ويعني هذا أن المعادلة الطبيعية الرابعة ليست مستقلة. بينما لدينا أربعة مقدرات للمعلمات ينبغي أن نجد لها حلاً $\hat{b}_0, \hat{b}_1, \hat{b}_2$ و \hat{b}_3 فإنه يتوافر لنا ثلاث معادلات مستقلة، فقط، وهكذا، فمن المستحيل تقدير هذه المعلمات.

(ب) بالتعويض عن Δi في معادلة الطلب وإعادة ترتيب الحدود، نحصل على:

$$M_{dt} = b_0 + b_1^* i_t + b_2^* i_{t-1} + u_t, \quad b_1^* = (b_1 + b_3) \text{ and } b_2^* = (b_2 - b_3).$$

٥- بأخذ التحويل اللوغاريتمي لدالة الإنتاج، نحصل على:

$$\log Q_t = \log A + \alpha_1 \log L_t + \alpha_2 \log (10,000 - L_t) + \alpha_3 \log K_t.$$

والآن، يمكننا استخدام الطريقة العادية لتقدير هذه المعادلة، طالما أن $\log L_t$ و

$\log (10,000 - L_t)$ ليسا مرتبطين ارتباطاً خطياً تاماً.

٦- نعلم من المتن أن :

$$\hat{b} = b + \frac{W_1 u_1}{A} + \dots + \frac{W_n u_n}{A}, \quad (1)$$

حيث إن $W_i = X_i - \bar{X}$ و $A = \sum (X_i - \bar{X})^2$ من (١) نحصل على:

$$(\hat{b} - b)^2 = \frac{W_1^2 u_1^2}{A^2} + \dots + \frac{W_n^2 u_n^2}{A^2} + \frac{W_1 W_2 u_1 u_2}{A^2} + \dots + W_i W_j u_i u_j + \dots$$

وهكذا

$$\sigma_b^2 = E(\hat{b} - b)^2 = \frac{\sigma_u^2}{\sum (X - \bar{X})^2} + E(\text{all cross product terms})$$

وبسبب الارتباط الذاتي، لم تعد القيمة المتوسطة لهذه الحدود الناتجة عن حاصل ضرب المتجهات صفرية، ومن ثم، لم تعد صيغة للتباين صحيحة. ٧- للتخلص من اختلاف التباين الموجود بالمعادلة، نقسم المعادلة على Y_i :

$$\frac{C_i}{Y_i} = b_0 \frac{1}{Y_i} + b_1 + b_2 \frac{A_i}{Y_i} + u_i^*,$$

حيث إن $u_i^* = u_i/Y_i$ ، وتكون المعادلات الطبيعية:

$$\sum \frac{C_i}{Y_i^2} = \hat{b}_0 \sum \frac{1}{Y_i^2} + \hat{b}_1 \sum \frac{1}{Y_i} + b_2 \sum \frac{A_i}{Y_i^2},$$

$$\sum \frac{C_i}{Y_i} = \hat{b}_0 \sum \frac{1}{Y_i} + n\hat{b}_1 + \hat{b}_2 \sum \frac{A_i}{Y_i},$$

$$\sum \frac{C_i A_i}{Y_i^2} = \hat{b}_0 \sum \frac{A_i}{Y_i^2} + \hat{b}_1 \sum \frac{A_i}{Y_i} + \hat{b}_2 \sum \frac{A_i^2}{Y_i^2}.$$

لاحظ أن هذه المعادلة لها حد ثابت ولذا نستخدم الشرط $\sum \hat{u}_i^2 = 0$.

٨- الخطوة الأولى: a_0 ، a_1 و a_2 بالطريقة العادية.

الخطوة الثانية: استخدم المعاملات المقدرة في الحصول على مجموعة من القيم

المقدرة للخطأ العشوائي، حيث:

$$\hat{\varepsilon}_i = Y_i - \hat{Y}_i = Y_i - \hat{a}_0 - \hat{a}_1 \Delta Y_i - \hat{a}_2 r_i.$$

الخطوة الثالثة: ادخل قيم $\hat{\varepsilon}_i$ في العلاقة:

$$\hat{\varepsilon}_i = \rho_1 \hat{\varepsilon}_{i-1} + \rho_2 \hat{\varepsilon}_{i-2} + u_i.$$

قدر P_2, P_1 عن طريق وضع :

$$\Sigma(\hat{u}_t \hat{\varepsilon}_{t-1}) = 0 \text{ و } \Sigma(\hat{u}_t \hat{t}_{t-2}) = 0$$

الخطوة الرابعة: حول النموذج الأصلي إلى :

$$I_t^* = a_0^* + a_1 \Delta Y_t^* + a_2 r_t^* + u_t,$$

حيث :

$$I_t^* = I_t - \hat{\rho}_1 I_{t-1} - \hat{\rho}_2 I_{t-2},$$

$$a_0^* = a_0 - \hat{\rho}_1 a_0 - \hat{\rho}_2 a_0,$$

$$\Delta Y_t^* = \Delta Y_t - \hat{\rho}_1 \Delta Y_{t-1} - \hat{\rho}_2 \Delta Y_{t-2},$$

$$r_t^* = r_t - \hat{\rho}_1 r_{t-1} - \hat{\rho}_2 r_{t-2}.$$

الخطوة الخامسة: قدر a_0^*, a_1 و a_2 بالطريقة المعتادة، ثم اجعل :

$$\hat{a}_0 = \hat{a}_0^* / (1 - \hat{\rho}_1 - \hat{\rho}_2).$$

الفصل السابع

١- (أ) المتغيرات الداخلية هي C_t, Y_t و I_t ، أما المتغيرات المحددة مسبقا فهي

$$C_{t-1}, Y_{t-1} \text{ و } r_t.$$

(ب) المشكلة الموجودة في المعادلة (1) هي أن Y_t مرتبطة بـ ε_{it} ، ولذلك، تكون

الطريقة هو إحلال \hat{Y}_t محل Y_t ، ونحصل على \hat{Y}_t عن طريق إجراء

انحدار لـ Y_t على جميع المتغيرات المحددة مسبقا وهي C_{t-1}, Y_{t-1} و r_t .

وهكذا سيكون \hat{Y}_t :

$$\hat{Y}_t = \hat{\gamma}_0 + \hat{\gamma}_1 C_{t-1} + \hat{\gamma}_2 Y_{t-1} + \hat{\gamma}_3 r_t.$$

ثم نقدر المعادلة (1)، حيث، بالطريقة المعتادة بعد أن نحل \hat{Y}_t محل Y_t ،

أي أن المعادلات الطبيعية حصل عليها عن طريق وضع المجاميع التالية

$$\Sigma(\hat{\varepsilon}_t^* Y_t) = 0 \text{ وأيضا } \Sigma(\hat{\varepsilon}_t^* C_{t-1}) = 0, \Sigma \hat{\varepsilon}_t^* = 0,$$

(ج) لتقدير المعادلة (3)، نستخدم الطريقة نفسها التي اتبعناها في تقدير (1). وسنحصل

على المعادلات الطبيعية عن طريق وضع المجاميع التالية مساوية للصفر.

$$\text{و : } \Sigma(\hat{\varepsilon}_2^* \hat{Y}_t) = 0 \text{ و } \Sigma(\hat{\varepsilon}_2^* r_t) = 0 \text{ و } \Sigma(\hat{\varepsilon}_2^* Y_{t-1}) = 0 \text{ و } \Sigma \hat{\varepsilon}_2^* = 0$$

٢- (أ) المعادلة الأولى مميزة حيث إن عدد المتغيرات المحددة مسبقاً في النموذج التي لا تظهر في المعادلة الأولى، \dot{M}_t ، أكبر من عدد المتغيرات الداخلية (\dot{P}_t) التي تظهر في المعادلة الأولى أو تساويه. ولكن ذلك ليس صحيحاً بالنسبة للمعادلة الثانية، ولذا تكون غير مميزة.

(ب) تكون طريقة التقدير للمعادلة الأولى على النحو التالي :

الخطوة الأولى: نحصل على (\hat{P}_t) عن طريق انحدار لـ (\dot{P}_t) على المتغيرات المحددة مسبقاً كافة وهي \dot{M}_t و UN_t وهكذا يكون (\hat{P}_t) هو :

$$\hat{P}_t = \hat{\gamma}_0 + \hat{\gamma}_1 \dot{M}_t + \gamma_2 UN_t$$

الخطوة الثانية: نجعل (\hat{P}_t) محل (\dot{P}_t) في المعادلة الأولى ثم نكمل كالعادة للحصول على معادلات المرحلة الثانية الطبيعية عن طريق وضع المجاميع التالية مساوية الصفر :

$$\sum \hat{\varepsilon}_{1t}^* = 0, \sum (\hat{\varepsilon}_{1t}^* UN_t) = 0, \sum (\hat{\varepsilon}_{1t}^* \hat{P}_t) = 0.$$

٣- (أ) للحصول على معادلات الشكل المختزل، نحل المعادلات (1) و (2) للحصول على L_t و W_t ويكون الشكل المختزل :

$$L_t = a_0^* + a_1^* P_t + a_2^* S_t + v_t,$$

$$W_t = b_0^* + b_1^* S_t + b_2^* P_t + \varepsilon_t,$$

$$a_0^* = \frac{a_0 + a_1 b_0}{1 - a_1 b_1}, \quad a_1^* = \frac{a_1 b_2}{1 - a_1 b_1}, \quad a_2^* = \frac{a_2}{1 - a_1 b_1},$$

$$v_t = \frac{a_1 u_{2t} + u_{1t}}{1 - a_1 b_1}, \quad b_0^* = \frac{b_0 + b_1 a_0}{1 - a_1 b_1}, \quad b_1^* = \frac{b_1 a_2}{1 - a_1 b_1},$$

$$b_2^* = \frac{b_2}{1 - a_1 b_1}, \quad \varepsilon_t = \frac{u_2 + b_1 u_{1t}}{1 - a_1 b_1}.$$

(ب) المشكلة الموجودة في المعادلة (1) هي أن W_t مرتبط بالخطأ العشوائي u_t . ولذلك نحل \hat{W}_t محل W_t ، ونحصل على \hat{W}_t عن طريق إجراء انحدار W_t على المتغيرات المحددة مسبقاً وهي S_t و P_t . لذلك فإن \hat{W}_t ستأخذ الشكل :

$$\hat{W}_t = \hat{\gamma}_0 + \hat{\gamma}_1 S_t + \hat{\gamma}_2 P_t.$$

بعدئذ، تقدر المعادلة (1) بالطريقة العادية بعد إحلال \hat{W}_t محل W_t أي أن المعادلات الطبيعية قد حصل عليها عن طريق مساواة المجاميع التالية بالصفر :

$$\sum \hat{u}_{1t}^* = 0, \sum (\hat{u}_{1t}^* \hat{W}_t) = 0, \sum (\hat{u}_{1t}^* S_t) = 0.$$

٤- (أ) إن إحدى الطرق البديهية لإدراك أن $D_{it(t-1)}$ يكون مرتبطاً مع u_{it} هي على النحو التالي : من المعادلة (1)، نجد أن $D_{it(t-1)}$ يعتمد على $u_{i(t-1)}$. ولكن، طالما أن $u_{it} = \rho u_{i(t-1)} + \varepsilon_{it}$ ، فإن u_{it} يعتمد، أيضاً على $u_{i(t-1)}$. وهكذا فإن u_{it} و $D_{it(t-1)}$ مرتبطان لأنهما يحتويان عنصراً مشتركاً.

(ب) إذا حلت المعادلة الأساسية حلاً متكرراً، فإنه يتبين لنا D_{it} تعتمد على ρ_t وعلى جميع قيمها المبطة وبصراحة يكون لدينا :

$$D_{it} = a_0 + a_0 a_2 + a_0 a_2^2 + \dots + a_1 P_t + a_1 a_2 P_{t-1} + a_1 a_2^2 P_{t-2} + \dots + u_{it} + a_2 u_{i(t-1)} + a_2^2 u_{i(t-2)} + \dots.$$

إضافة إلى ذلك، نجد أن D_{it} سوف تعتمد على u_{it} وعلى قيمها المبطة. وهكذا فإن D_{it} سيتماد، فقط، على ρ_t وعلى جميع قيمها المبطة. فإذا أخذنا في الحسبان المعادلة السابقة التي تربط D_{it} بقيم كل من ρ_t و u_t بوصفها معادلة ذات شكل مختزل فسيمكننا أن ننفذ طريقة م ص م الموصوفة في الكتاب لنموذج تكون به جميع المتغيرات المحددة مسبقاً إما غير معلومة أو لا تتوافر لدينا مشاهدات عنها. وبمعنى آخر، يمكننا تنفيذ طريقة م ص م عن طريق إجراء انحدار D_{it} على P_t و $\hat{D}_{i(t-1)}$ ، حيث أن $\hat{D}_{i(t-1)}$ تحسب عن طريق

انحدار D_{it-1} على P_t وبعض قيمه المبطة، ثلاثة منها، مثلاً.
 ٥- بالتعويض من (1) و (2) نحصل على

$$X_{2t} = b_0^* + b_1^* X_{1t} + v_t^*, \quad (1')$$

حيث

$$b_0^* = \frac{c_0 + c_1 b_0}{1 - c_1 b_2}, \quad b_1^* = \frac{c_1 b_1}{1 - c_1 b_2}, \quad v_t^* = \frac{c_1 u_{1t} + u_{2t}}{1 - c_1 b_2}.$$

وبضرب (1') في u_{1t} وأخذ القيم المتوقعة نحصل على :

$$E(X_{2t} u_{1t}) = b_0^* E(u_{1t}) + b_1^* E(u_{1t} X_{1t}) + \frac{c_1 E(u_{1t}^2)}{1 - c_1 b_2} + \frac{E(u_{1t} u_{2t})}{1 - c_1 b_2}$$

وبتذكر أن :

$$E(u_{1t}) = 0, \quad E(u_{1t} X_{1t}) = 0, \quad E(u_{1t} u_{2t}) = \text{cov}(u_1, u_2),$$

يكون لدينا :

$$E(X_{2t} u_{1t}) = \frac{c_1 \sigma_1^2}{1 - c_1 b_2} + \frac{\text{cov}(u_1, u_2)}{1 - c_1 b_2} \neq 0$$

إلا إذا كان :

$$\frac{c_1 \sigma_1^2}{1 - c_1 b_2} = - \frac{\text{cov}(u_1, u_2)}{1 - c_1 b_2}$$

٦- (أ) لإثبات أن طريقة م ص م تفشل في تقدير المعادلة الأولى، تتبع الخطوات

التالية: نجري أولاً انحدار \hat{P}_t على جميع المتغيرات المحددة مسبقاً في

النموذج ويعطينا هذا :

$$\hat{P}_t = \hat{c}_0 + \hat{c}_1 (UN_t).$$

وبإحلال \hat{P}_t محل \hat{P}_t في المعادلة الأولى، نحصل على :

$$\hat{W}_t = a_0 + a_1 \hat{P}_t + a_2 (UN_t) + \varepsilon_{1t}^*.$$

نلاحظ أنه طالما أن \hat{P}_t و UN_t مرتبطين ارتباطاً تاماً، فإنه سيكون من

المستحيل تقدير a_1 و a_2 وهكذا تفشل طريقة م ص م إذا حاولنا تقدير

المعادلة الأولى.

(ب) لا تفشل طريقة م ص م في تقدير المعادلة الثانية، لأننا هنا لانواجه هنا بمشكلة الارتباط الخطي المتعدد. نحصل أولاً على \hat{W}_t عن طريق عمل انحدار لها على UN_t وبعدئذ، نحل \hat{W}_t محل \hat{W}_t في المعادلة الثانية ونكمل بالطريقة المعتادة لاشتقاق المعادلات الطبيعية عن طريق مساواة المجاميع التالية بالصفر:

$$\sum \hat{\varepsilon}_{2t}^* = 0, \sum (\hat{\varepsilon}_{2t}^* \hat{W}_t) = 0$$

وتكون المعادلات الطبيعية هي :

$$\sum \dot{P}_t = n\hat{b}_0 + \hat{b}_1 \sum \hat{W}_t$$

$$\sum (\dot{P}_t \hat{W}_t) = \hat{b}_0 \sum \hat{W}_t + \hat{b}_1 \sum \hat{W}_t^2.$$

هذه المعادلات مستقلة خطياً ولذا يمكننا حلها للحصول على \hat{b}_0 و \hat{b}_1 .
٧- (أ) المعادلة مميزة حيث إن عدد المتغيرات المحددة مسبقاً في نظام المعادلات التي لا تظهر في المعادلة موضع الاهتمام أكبر من عدد المتغيرات المستقلة الداخلية.

(ب) لا يمكننا تقدير المعادلة باستخدام م ص م لأننا نحتاج لمتغيرين محددين مسبقاً في الأقل من تلك التي لا تظهر في المعادلة، ولكن طالما أنه تتوافر لدينا مشاهدات عن واحد، فقط، من هذه المتغيرات (X_{2t}) فإنه يمكن إثبات أنه في ظل البيانات القاصرة لا توجد طريقة تمكننا من تقدير معادلة الانحدار في هذه المسألة.

٨- (أ) كلا المعادلتين مميزتان، طالما أن كلاهما يحتوي على عدد من المتغيرات المحددة مسبقاً والمستبعدة من المعادلة مساوياً لعدد المتغيرات المستقلة الداخلية. لاحظ أن X_t و X_t^2 ليسا مرتبطين خطياً ارتباطاً تاماً. ولذلك، يمكن اعتبار X_t^2 متغيراً محدداً مسبقاً وغير مشتمل عليه في المعادلة الثانية.

(ب) للحصول على الشكل المختزل، نعوض عن Y_{2t} في المعادلة الأولى وعن Y_{1t} في المعادلة الثانية وبعد إعادة ترتيب الحدود يكون لدينا :

$$Y_{1t} = a_1^* + b_1^* X_t^2 + c_1^* X_t + v_{1t}^*,$$

$$Y_{2t} = a_2^* + b_2^* X_t^2 + c_2^* X_t + v_{2t}^*,$$

حيث :

$$a_1^* = \frac{a_1 + c_1 a_2}{1 - c_1 c_2},$$

$$b_1^* = \frac{b_1}{1 - c_1 c_2},$$

$$c_1^* = \frac{c_1 b_2}{1 - c_1 c_2},$$

$$v_{1t}^* = \frac{c_1 \varepsilon_{2t} + \varepsilon_{1t}}{1 - c_1 c_2},$$

$$a_2^* = \frac{a_2 + c_2 a_1}{1 - c_1 c_2},$$

$$b_2^* = \frac{c_2 b_1}{1 - c_1 c_2},$$

$$c_2^* = \frac{b_2}{1 - c_1 c_2},$$

$$v_{2t}^* = \frac{c_2 \varepsilon_{1t} + \varepsilon_{2t}}{1 - c_1 c_2}.$$

(ج) نستخدم طريقة م ص م في تقدير المعادلة الأولى ، وهكذا نكون انحدارا لـ Y_{2t} على المتغيرات المحددة مسبقا كافة للحصول على :

$$\hat{Y}_{2t} = \hat{a}_2^* + \hat{b}_2^* X_t^2 + \hat{c}_2^* X_t.$$

بعد ذلك ، نحل \hat{Y}_{2t} محل Y_{2t} في المعادلة الأولى ونكمل المنهج لتقدير المعادلة بالطريقة المعتادة ، أي نشق المعادلة الطبيعية بوساطة المجاميع

التالية بالصفر:

$$\sum \hat{\varepsilon}_{it}^* = 0, \quad \sum (\hat{\varepsilon}_t^* X_t^2) = 0, \quad \sum (\hat{\varepsilon}_{it}^* \hat{Y}_{2t}) = 0$$

٩- (أ) المعادلة الأولى غير مميزة، بسبب أن t سيكون ثابتا في التحليل المقطعي. وطالما أنه يوجد لدينا ثابت في المعادلة الأولى، فإننا لا يمكن أن نستفيد منه بوصفه متغيرا محددا مسبقا مستبعدا للحصول على r_{it} . المعادلة الثانية مميزة بسبب استبعاد متغير المبيعات. افترض الآن أنه يتوافر لدينا بيانات سلسلة زمنية لعدد T من الفترات لمتغيرات نموذجنا. افترض، أيضا، أن هذه المنشآت التي عددها N تمثل قسما صغيرا من الاقتصاد القومي. ولذا، يمكن اعتبار r_t متغيرا خارجيا، حيث، فإن المعادلة الأولى ستكون مميزة، لأن r_t في هذه الحالة يمكن اعتباره متغيرا محددا مسبقا مستبعدا. في هذه الحال، ستقدر معادلة الاستثمار على النحو التالي: أولا: باستخدام مشاهدات سلسلة زمنية عددها T ، نكون انحدارا لكل من r_{it} على r_t وعلى متغير المبيعات المناظر للحصول على:

$$\hat{r}_{it} = \hat{\gamma}_{0i} + \hat{\gamma}_{1i} S_{i(i-1)} + \hat{\gamma}_{2i} r_t, \quad i=1, \dots, N.$$

والآن أحل \hat{r}_{it} محل r_{it} وأكمل لتقدير المعادلة الأولى بالمنهج المعتاد. وعند عملنا ذلك لاحظ أنه ينبغي أن يتوافر لدينا NT مشاهدات في المرحلة الثانية.

(ب) نحصل على الشكل المختزل لـ I_{it} عن طريق التعويض عن r_{it} في المعادلة الأولى، وبعد إعادة ترتيب الحدود، نحصل على:

$$I_{it} = a^* + b_1^* r_t + b_2^* S_{it-1} + v_{it}^*,$$

حيث:

$$a^* = \frac{a}{1-b_1 b_3}, \quad b_1^* = \frac{b_1}{1-b_1 b_3},$$

$$b_2^* = \frac{b_2}{1-b_1 b_3}, \quad v_{it}^* = \frac{b_1 \varepsilon_{it} + u_{it}}{1-b_1 b_3}$$

ثبت المصطلحات

أولاً: عربي - إنجليزي



Lag	إبطاء
Almon lag	آلمون
Koyck lag	كويك
Time trend	اتجاه زمني
Consistency	اتساق
Probability	احتمال
Joint probability	مشترك
Statistic	إحصائية
Goldfeld-Quandt test (G-Q)	اختبار جولدفيلد كوندات (ج - ك)
One tailed test	الذيل الواحد
Two tailed test	الذيلين
Hypotheses testing	الفرضيات
Heteroscedasticity	اختلاف التباين
Correlation	ارتباط
Autocorrelation	ذاتي
Independence	استقلال

Assumption	افتراض
Econometrics	اقتصاد قياسي
Regression	انحدار
Standard deviation	إنحراف معياري
ب	
Residual	باقي
ت	
Formalization	تأطير
Trade off	تبادل
Variance	تباين
Specification of Model	تحديد النموذج
Semilog transformation	تحويل شبه لوغاريتمي
Reciprocal transformation	عكسي
Bias	تحيز
Multicollinearity	تعدد العلاقات الخطية
Covariance	تغاير
Estimate	تقدير
Point estimate	النقطة
Proxy	تقريبي
Identification	تمييز
Forecast	تنبؤ
Normal distribution	توزيع طبيعي
Sampling distribution	المعينة
Fit	توفيق

Expectation	توقعات
Combination	توليفة
	
Additive constant	ثابت تجميعي
Homoscedasticity	ثبات التباين
	
Goodness of fit	جودة التوفيق
	
Univariate case	حال المتغير الواحد
Lower bound	حد أدنى
Error term	الخطأ
	
Exogenous	خارجي
Fallacy of composition	خدعة التجميع
Disturbance term	خطأ عشوائي (حد الخطأ)
Standard error	معياري
Type I error	من النوع الأول
Type II error	من النوع الثاني
Linear	خطي
	
Endogenous	داخلي
Function	دالة
Subscript	دليل سفلي

Formally	ر	رياضيا، اصطلاحيا
Spurious	ز	زائف
Time series	س	سلسلة زمنية
Scatter diagram	نق	شكل انتشار
Reduced form		مختزل
Explicit	صر	صريح
Endpoint	ط	طرفي
Dependence	ع	عدم استقلال ، اعتماد
Sorting		عزل ، فصل
Random		عشوائي
Scale factors		عوامل ترجيح
Unbiased	غ	غير متحيز
Distributed lag	فا	فترة إبطاء

Alternative hypothesis	فرضية بديلة
Null hypothesis	العدم
Discrepancy	فرق
	
Density	كثافة
	
Continuous	متصل
Polynomial	متعدد الحدود
Variables	متغيرات
Dummy variable	متغير صوري
Explanatory variable	مفسر
Overall mean	متوسط حسابي شامل
Mean square error	مربع الخطأ
Population	مجتمع
Subset	مجموعة جزئية
Error sum of squares (ESS)	مجموع مربعات الأخطاء
Regression sum of squares (RSS)	الانحدار
Predetermined	محدد مسبقاً
Bounded	محدود
Two stage least squares (TSLS)	مربعات صغرى ذات مرحلتين (م ص ع)
Ordinary least squares (OLS)	عادية (م ص ع)
Instrumental	مساعدة
Significance level	مستوى المعنوية

Monotonic	مضطرد
First difference equations	معادلات الفروق من الدرجة الأولى
Structural equations	هيكلية
Under- identified equation	معادلة ناقصة التمييز
Operator	معامل
Coefficients	معاملات
Coefficient of Determination (R^2)	معامل التحديد
Adjusted R^2	المعدل
Reasonable	معقولة
Parmeter	معلمة
Limited-information	معلومات محدودة
Offsetting	معوضة
Estimators	مقدرات
Cross - sectional	مقطعي
Deflated	مكمش
Degenerated	منحل
Phillips curve	منحنى فليبس
Critical region	منطقة حرجة
Rejection region	الرفض
Acceptance region	القبول
Circluar reasoning	منطق دائري
Discrete	منقطع
Exact	مؤكد، يقيني
Marginal propensity to consume (MPC)	ميل حدي للاستهلاك



Corollaries	نتائج تابعة
Central tendency	نزعة مركزية
Over determined	نظام زائد التحديد
Inflection points	نقاط انقلاب
Typical	نمطي
Autoregressive Model	نموذج للإنحدار الذاتي
Finite	نهائية
Open ended	نهاية مفتوحة



Mean	وسط حسابي
Events	وقائع

ثانياً : (انجليزي - عربي)

A

Acceptance region	منطقة القبول
Additive constant	ثابت تجميعي
Adjusted R^2	معامل التحديد المعدل
Almon lag	ابطاء المون
Alternative hypothesis	فرضية بديلة
Assumption	افتراض
Autocorrelation	الارتباط الذاتي
Autoregressive Model	نموذج للإنحدار الذاتي
Average out	توزيع

B

Bias	تحيز
Bounded	محدودة

C

Central tendency	نزعة مركزية
Circular reasoning	منطق دائري
Coefficient of Determination	معامل التحديد
Coefficients	معاملات
Combination	توليفة
Consistency	اتساق
Continuous	متصل
Corollaries	نتائج تابعة

Correlation	ارتباط
Covariance	تغاير
Critical region	منطقة حرجة
Cross - sectional	مقطعي

D

Deflated	مكمش
Degenerated	منحل
Density	كثافة
Dependence	عدم استقلال (اعتماد)
Determination coefficient	معامل التحديد
Discrepancy	فرق
Discrete	منقطع
Distributed lag	فترة ابطاء
Disturbance term	خطأ عشوائي (حد الخطأ)
Dummy variable	متغير صوري

E

Econometrics	اقتصاد قياسي
Endogenous	داخلي
Endpoint	طرفي
Error sum of squares (ESS)	مجموع مربعات الاخطاء
term	حد الخطأ (خطأ عشوائي)
Estimate	تقدير
Estimators	مقدرات
Events	وقائع
Exact	مؤكد يقيني

Exogenous	خارجي
Expectation	توقعات
Explanatory variable	متغير مفسر
Explicit	صريح

F

Fallacy of Composition	خدعة التجميع
Finite	محدودة
First difference equations	معادلات الفروق من الدرجة الأولى
Fit	توفيق
Forecast	تنبؤ
Formalization	تأطير
Formally	رياضيا، اصطلاحيا
Function	دالة

G

Goldfeld-Quandt test (G-Q)	اختبار جولدفيلد كوندات (ج - ك)
Goodness of fit	جودة التوفيق

H

Heteroscedasticity	اختلاف التباين
Homoscedasticity	ثبات التباين
Hypotheses testing	اختبار الفرضيات

I

Identification	تمييز
Independence	استقلال
Inflection points	نقاط انقلاب
Instrumental	مساعد

Joint probability	J	احتمال مشترك
Koyck lag	K	إبطاء كويك
Lag	L	إبطاء
Limited-information		معلومات محدودة
Linear		خطي
Lower bound		حد أدنى
Marginal propensity to consume (MPC)	M	الميل الحدي للاستهلاك (م ح س)
Mean		وسط حسابي
square error		متوسط مربع الخطأ
Monotonic		مضطرد
Multicollinearity		تعدد العلاقات الخطية
Normal distribution	N	توزيع طبيعي
Null hypothesis		فرضية العدم
Offsetting	O	معوضة
One tailed test		اختبار الذيل الواحد
Open ended		ذو نهاية مفتوحة
Operator		معامل
Ordinary least squares (OLS)		طريقة المربعات الصغرى العادية (م ص ع)

Overall mean	متوسط شامل
Over determined system	نظام زائد التحديد

P

Parmeter	معلمة
Perdiction	تنبؤ
Phillips curve	منحنى فليبس
Point estimate	تقدير النقطة
Polynomial	متعدد الحدود
Population	مجتمع
Predetermined	محدد مسبقا
Probability	احتمال
Proxy	تقريبي

R

Random	عشوائي
Reasonable	معقولة
Reciprocal transformation	تحويل عكسي
Reduced form	شكل مختزل
Regression	انحدار
sum of squares (RSS)	مجموع مربعات الانحدار
Rejection region	منطقة الرفض
Relabeling	إعادة صياغة
Residual	باقي

S

Sampling distribution	توزيع المعاينة
Scale factors	عوامل ترجيح

Scatter diagram

شكل الانتشار

Semilog transformation

تحويل شبه لوغاريتمية

Significance level

مستوى المعنوية

Sorting

عزل (فصل)

Specification of Model

تحديد النموذج

Spurious

زائف

Standard deviation

انحراف معياري

error

خطأ معياري

Statistic

إحصائية

Structural equations

معادلات هيكلية

Subscript

دليل سفلي

Subset

مجموعة جزئية

T

Time series

سلسلة زمنية

trend

اتجاه زمني

Total sum of squares (TSS)

المجموع الكلي للمربعات

Trade-off

تبادل

Two stage least squares (2SLS)

المربعات الصغرى ذات المرحلتين (م ص م)

tailed test

اختبار الذيلين

Type I error

الخطأ من النوع الأول

II error

الخطأ من النوع الثاني

Typical

نمطي

U

Unbiased

غير متحيز

Under-identified equation

معادلة ناقصة التميز

Univariate case

حال المتغير الواحد



Variables

متغيرات

Variance

تباين

كشاف الموضوعات



ابطاء آلون ٢٥٢، ٢٦٢،

كويك ٢٤٧، ٢٥١

اتساق ٢٧، ٢٨، ٤٩، ٥٠، ٣٨٥،

٣٩٠، ٣٩٣، ٤٠٣، ٤٠٤

احتمال مشترك ٣٤، ٣٧

إحصائية ١٥١-، ١٥٤، ٤٩٣

اختبار جولد فيلد-كوندات ٣٥٤، ٣٦٠

الذيل الواحد ١٥٣

الذيلين ١٥١، ١٥٣

اختلاف التباين ٣٣٨، ٣٦٦

ارتباط ٥٢، ٦٥، ١١٦، ١١٨

ذاتي ٣١٠، ٣٣٨، ٤٣١، ٤٣٦

استقلال ١٨، ١٩، ٣٩، ٤٠، ٤٤، ٥٢

انحراف معياري ٢٠



باقي ٣٠٥

بيانات مقطعية ٨٦



تباين ٩٥، ١٠٥، ١٢١، ١٢٧، ٢٣٨،

٣٠٤، ٣٠٦

تحديد النموذج ٣٩٣، ٣٩٦

تحويل شبه لوغاريتمي ١٦٤، ١٦٨

عكسي ١٥٤، ١٦٠

لوغاريتمي ١٦٠، ١٦٣، ٢٧٣، ٢٧٦

تحيز ٢٥، ٢٦، ٤٦، ٤٩، ٩٢، ٩٥، ٤١٧

تعدد العلاقات الخطية ٢٠٣، ٢٠٩، ٢١١،

٣٠٢، ٣١٠

تغاير ٣٤، ٤١، ٤٩

تقدير النقطة ١٣٢

تنبؤ ١٨٣، ١٩١، ٣٠٩، ٣١٠

توزيع طبيعي ١٣٤، ١٣٧، ٤٩٢،

٤٩٤F-٤٩٧

المعينة ٤٧

توقعات ١٩، ٢٣



ثبات التباين ٣٣٨



خطأ عشوائي ٦٧، ٧١، ٧٣، ٧٦، ١٣٣،

١٣٤، ٢٠٢، ٢٠٣

معياري ١٥١، ١٥٤

من النوع الأول ١٣٨، ١٣٩

خطأ من النوع الثاني ١٣٨ ، ١٣٩

د

دالة الكثافة المشتركة ٢٨ ، ٣٧

ش

شكل انتشار ٥ ، ٤٢ ، ٤٣

دالي ١٥٤ ، ١٧٦ ، ٢٧٣ ، ٢٨٥

مختزل ٣٩٨ ، ٤٠٠

ع

عدم التحيز ٢٥ ، ٢٦ ، ٤٦ ، ٤٩ ، ٩٢ ،

٩٥ ، ٢٣٧

علاقة زائفة ٨

ف

فترات ثقة ١٣١ ، ١٥٠ ، ١٨٤ ، ١٩١ ،

٢١٦ ، ٢١٩ ، ٣٠٤

ك

كثافة ١٦ ، ٢٨ ، ٣٨

م

متعدد الحدود ٢٧٧ ، ٢٩٢

متغيرات مبطأة ١٧٧ ، ١٨٣ ، ٢٤٣ ، ٢٦٢ ،

٣٩٧ ، ٣٩٨

متغير صوري ٢٦٢ ، ٢٧٣

عشوائي ١٥ ، ١٧

مساعد ٧٧ ، ٨٥ ، ٢٠٥ ، ٢٠٩

متوسط مربع الخطأ ٣٢٠

مجموع مربعات الأخطاء ١١٣ ، ١١٤

الانحدار ١١٣ ، ١١٤

مربعات صغرى ذات مرحلتين ٣٨٤ ، ٣٩٣ ،

٤٠٠ ، ٤٠٨ ، ٤٢٠ ، ٤٣١

عادية ١٠٦

مشكلة التمييز ٤٠٨ ، ٤٢٠ ، ٤٤٩ ، ٤٦٤

معادلات طبيعية ٨٠ ، ٨٣ ، ٢٠٦ ، ٢٠٩

هيكلية ٣٩٨ ، ٤٠٠

معامل التحديد ١١٠ ، ١١٥ ، ٢١٩ ، ٢٢٧

المعدل ٢٢٤ ، ٢٢٧

مقدرات ٢٥ ، ٢٨ ، ٥٨ ، ٦١ ، ٧٧ ، ٨٥

١٠٣ ، ١٠٥ ، ٢٠٥ ، ٢٠٩ ،

٢٣٤ ، ٢٣٨

مناطق القبول والرفض ١٤٠

منحني فليس ٤٣ ، ١٥٤ ، ١٦٠ ، ٢٨٣ ،

٢٨٤

منطقة حرجة ١٤٠

ن

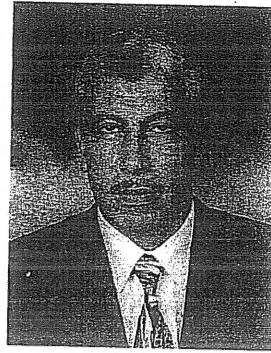
نظام زائد التحديد ٤٤٩

نموذج الانحدار ٧١ ، ٧٧ ، ٢٠٢ ، ٢٠٥



د. عبدالقادر محمد عبدالقادر عطية

- حصل على درجة الدكتوراه في الاقتصاد من جامعة نوتردام بالولايات المتحدة الأمريكية عام ١٩٨٧ م.
- يشغل الآن وظيفة أستاذ الاقتصاد بكلية التجارة، جامعة الإسكندرية.
- أعير إلى جامعة الملك سعود، فرع القصيم لمدة خمس سنوات خلال الفترة ١٤١٠-١٤١٥ هـ.
- صدر له عدة كتب منها: طرق قياس العلاقات الاقتصادية مع تطبيقات على الحاسب الإلكتروني، دراسات الجدوى التجارية والاقتصادية والاجتماعية مع تطبيقات على الحاسب الإلكتروني، الاقتصاد الصناعي بين النظرية والتطبيق، اقتصاد المملكة العربية السعودية ونظرة تحليلية (مشترك).
- نشر العديد من البحوث في مجالات: البطالة، توظيف الأموال، سوق الأسهم، اقتصاديات المخدرات، واقتصاديات الغش والتطفيف، اقتصاديات السلع الاستراتيجية والحرب الاقتصادية وسياسة التسعير في الواقع.



د. المرسي السيد أحمد حجازي

- حصل على درجة الدكتوراه في الاقتصاد من جامعة كونتكت بالولايات المتحدة الأمريكية عام ١٩٨٥ م.
- يشغل الآن وظيفة أستاذ الاقتصاد العام ورئيس قسم المالية العامة بكلية التجارة، جامعة الإسكندرية.
- أعير إلى جامعة الملك سعود، فرع القصيم، لمدة ست سنوات خلال الفترة ١٤١١-١٤١٧ هـ.
- صدر له عدة كتب: اقتصاديات الخدمات العامة، مبادئ الاقتصاد العام، ضرائب الدخل والثروة والإنفاق في لبنان، النظم الضريبية بين النظرية والتطبيق كما صدر له بالإنجليزية Tax Systems in Practice.
- نشر العديد من البحوث في مجالات: تقويم النظم الضريبية والسياسات المالية، الإنفاق العام والزكاة، الرفاهة الاقتصادية، المشروعات العامة، اقتصاديات الموارد البيئية وتسعير المياه، والتكاليف الاقتصادية لتلوث البيئة.

